

**FIGURE 1**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG  
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT  
TTCAGCAAGGCCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT  
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG  
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG  
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA  
TTAAACAAGGCTGGCCTTGCTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCT  
CCAGAAAACAACCCCTTTTGTGTCACATGTAAGTGAGCTGTGCTTACCTTTGGTATGGGCTCAT  
TATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC  
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC  
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG  
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT  
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA  
TGGATTAACCCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA  
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACCTTAATCAAGGCTGACAGTAACACT  
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT  
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT  
ATG

## **FIGURE 2**

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIYSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV  
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG  
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW  
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR  
LLSRDI

### **Important features:**

#### **Type II transmembrane domain:**

amino acids 13-33

#### **Other Transmembrane domains:**

amino acids 54-73, 94-113, 160-180, 122-141

#### **N-myristoylation sites.**

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

**FIGURE 3**

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC  
 GTGTGAGGGGGCCTGTGGCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC  
 CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG  
 GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG  
 GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG  
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG  
 TTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACACGA  
 CTGCTTTTTTCTGCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC  
 CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTGGTGTGATTGGAG  
 TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCCATACACTTACATGTCTTACTTC  
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGTGCAAACCATGGATATGAT  
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA  
 AACCATCAGGTTTCTGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAATCTTACT  
 CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAACAGCTGATCT  
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTGGTT  
 ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTGATCGAGTT  
 GGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT  
 GAAGTTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT  
 TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCTTG  
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT  
 AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACC GTTGGTTTG  
 ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA  
 GAGAAGCAAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAAATTTA  
 GATATAAGAGGGGGGAAAAATGGAACCAGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC  
 ATTTTTACCTTCATAGCATACTCCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG  
 AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG  
 TGTAGAGGCGGAGAGGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT  
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTTACATGGAAAAGGTTATAGCTTG  
 CCTTGAGATTGACTCATTAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAGGGCGGCCGCG  
 ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

## **FIGURE 4**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV  
LNSSSRYPFWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDP  
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERLLQ  
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRQ  
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVPVTRGIEI  
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY  
FVSSVLLIRMSMPLEYRTIITEVLGELQNFYHRWFDVIFLVSAISSILFLYLAHKQAPEKQMAP

### **Important features:**

#### **Signal peptide:**

amino acids 1-23

#### **Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,  
425-444

#### **N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

#### **Eukaryotic cobalamin-binding proteins**

amino acids 151-160

**FIGURE 5**

AGCAGGGAAATCCGGATGTCFCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC  
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCCTCATATCACCAGTGGCCATCTGAGGT  
 GTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTTCCTCTCACT  
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT  
 GTCATGCAGAAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG  
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC  
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTGGAGATGGATTGCTGGTCACTCTAG  
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTGGAAGTTCCAGTGA  
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA  
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAACGCAACACAACAACAGAATTTATTGT  
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTACTCTACAATACCTGCCCTACTACTACTCCTC  
 CTGTCAGCTTCCACTTCTATTCCACGGAGAAAAAAATGATTTGTGTACAGAAGTTTATTATG  
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA  
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGTCTGCTCTCTCTCTTTGGTGCTG  
 CAGCTGGTCTTGGATTTTGCTATGTCAAAAGGTATGTGAAGGCCCTCCCTTTTACAACAAGAAT  
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCTAA  
 TGAGGAATCAAAGAAAACCTGATAAAAACCCGAAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC  
 GATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCA  
 CCCTTGGTTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT  
 GCCCTTCTCCTTATGTAAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCCT  
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGCAAAGTGAAGGAC  
 CTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTTG  
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCTTTCTTCA  
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCTTCTGAGCCCGTAAGAGCAAAAGAAT  
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA  
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAC  
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA  
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA  
 AAAATAAAAACTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT  
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC  
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATGACATAGTAG  
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT  
 TTTTATTTTGTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT  
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC  
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT  
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

**FIGURE 6**

MARCFSLVLLLSIWTTTRLLVQGSRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG  
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN  
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTTPPAPASTSIPR  
RKKLICVTEVFMETSTMSTETETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK  
RYVKAFFFTNKNQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

**Signal sequence:**

amino acids 1-16

**Transmembrane domain:**

amino acids 235-254

**N-glycosylation site.**

amino acids 53-57, 130-134, 289-293

**Casein kinase II phosphorylation site.**

amino acids 145-149, 214-218

**Tyrosine kinase phosphorylation site.**

amino acids 79-88

**N-myristoylation site.**

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

**FIGURE 7**

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCGG  
GGCTCCCGGCGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGGCGG  
CGGCTGCGGGCGCAGAGCGGAGATCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG  
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGTCCAGTCAAGCCCGGCCG  
GCTCTCAGTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTATGAT  
GGAGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA  
AAGCATCATCAGAAGTGAACCTGGCAAACCTACCTCCCAGCTATCACAATGAGACCAACACAGAC  
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACCAGAC  
TGGACAAATGGTCTTTTCAGAGACAGTTATCAGATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC  
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC  
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGTGTGGAGACCA  
CTGTGTGTCTGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGAGCCATCTGTGACA  
ACCAGAGGGAAGTGCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGTGTTCCCTGTGTGC  
ACACCCCTGCCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC  
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCTCCTCTGCCAGC  
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGG  
GAGATCCTGCTGCCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG  
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG  
CCGCCGTGCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA  
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA  
CATCTTCTTCCCAGTAAGTTTCCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTACAGT  
CCCCCAGGCTGTTCTCCAGGCTTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAAACTGCA  
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCCTTACCAGTTGGCAGACAGCCG  
TTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC  
TGATTGGTTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAATG  
CAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC  
AGATGAAATGTTCTGTTTACCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC  
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG  
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC  
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT  
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT  
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAAGTAAA  
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT  
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCA  
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC  
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC  
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG  
TTTTCAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTTAAAGTTGCA  
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAA  
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

**SECRET**

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

App\_ID=10063525



## FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC  
 GGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG  
 GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT  
 AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCCT  
 CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC  
 AACTGGGTACTGGGCATTTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCCGGTTCATCTTG  
 GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCGCTGCATCATGTGTGTTT  
 CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA  
 TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAATGCGTTCATGCTACTCATGCGAAAC  
 ATTGTCAGGGTGGTCTGCTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT  
 GGTCGGAGGCGTGGGGTCTGTCTTCTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG  
 ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT  
 GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCTCTGCTTCTCT  
 GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA  
 AGATTCTGGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG  
 CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT  
 CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT  
 TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG  
 AAACCTCCGTCTCTATTAAAAATACAAAATTAGCCGAGAGTGGTGGCATGCACCTGTATCCCA  
 GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA  
 GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA  
 AAGATTTTATTAAAGATATTTTGTTAACTC

## **FIGURE 10**

RTRGRTRGGCEKVPINTSCNPHTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGFWTL  
 NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL  
 EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN  
 IVRVVLDKVTDLLLFFGKLLVVGGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY  
 VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

### **Important features:**

#### **Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

#### **N-glycosylation sites.**

amino acids 16-20, 27-31, 289-293

#### **Hypothetical YBR002c family proteins.**

amino acids 276-288

#### **Ammonium transporters proteins.**

amino acids 204-231

#### **N-myristoylation sites.**

amino acids 60-66, 78-84

#### **Amidation site.**

amino acids 306-310

### **FIGURE 11**

GCCCCGCGCCCCGGCGCGGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC  
TCCCTGCTCAGCTGCGCGTCTTGCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCGCG  
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCAGTTCTTCTCTTCTGCGGGGTGCTGGTGTCCA  
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC  
GGGATCCCCACCGTCTTGAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG  
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCA  
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCTTGATCCTGGTGGGCCTC  
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTACCACATCTGGTTCTACTTCGGCGTCGTGGG  
CTCCTTCTCTTTCATCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGT  
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC  
TTTCTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCAGCGCTGCCA  
CGAGGGCAAGGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTGCG  
CCAAGGTCCAGGACGCCACGCCAACTCGGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCATG  
TTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAACCCAGCT  
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG  
TGGGCCTCATCATCTTCTCTCTGTGCAACCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG  
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGT  
GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGTACTCCTTCTTCC  
ACTTCTGCCTGGTGTGCGCTCACTGCACGTATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG  
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT  
CCTCTACCTGTGGACCTTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA  
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC  
CCCACACCAATCAGCCAGGTGAGCCCCCACCCTGCCCGAGCTCCAGGACCTGCCCCCTGAGCCGGGC  
CTTCTAGTCGTAGTGCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC  
ACCCACACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCCTGTTGCCCATACTCAGCATCTCGGATGAA  
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCACAG  
TGGGCGATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGACCTGCCCCCTTCTCTG  
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTTCTAATAAACAAGCCAGTGCCTGTAAAAAAA

## FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLI FTFFLFLGVLVSIIMLSPGVESQL  
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAA FFFFFFFLMLCVSSSRDPRAAIQ  
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE  
ECDSRAWYAGLFFFTLLFYLLSIAAVALMEMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV  
QDAQPN SGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI  
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY  
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICASWAGLLLYLWTLVAPLLLRNRD  
FS

### Signal sequence:

amino acids 1-20

### Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

**FIGURE 13**

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA  
 GGTGGAAGAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG  
 CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA  
 CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC  
 TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG  
 TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT  
 TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGTCATATGCTGTGTGCAG  
 ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG  
 TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTTCATC  
 CTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA  
 CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG  
 GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT  
 GAGAAACCACTTTTAGAACTATTGAGTACTACTTTTGTTAAATGTGAAAAACCTCACAGAAAGTC  
 ATCGAGGCAAAAAGAGGCAGGCAGTGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC  
 CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC  
 CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTATCCTCTCT  
 TCAGTGAGACTGAGCCTGATGTGTTAACAATAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATC  
 AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC  
 AGAATGGAATTTTTTGTGTTTCATGTCTCAGATTTATTTGTATTTCTTTTTAACACTCTACATT  
 TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTG  
 ACATGTCAATGTGGCTAGTTTTATTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA  
 CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT  
 TTATCATGAAATCATGTTTTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC  
 ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTTCAGCCAATTTCAAATAAAATGAA  
 CTAAATTAAAAA

## **FIGURE 14**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF  
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT  
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER  
AALIPGGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-20

#### **Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

#### **N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

**FIGURE 15**

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG  
 CCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAGATGTGCTCCAG  
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGGGCTGCCAT  
 CCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC  
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC  
 AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACCAGATGCCAGCC  
 TGCCACGCGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG  
 CATGAAATCACC AATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTTACCTGGGCAAGAA  
 CCGCATCCGCCACATCCAGCCTGGTGCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC  
 AGGACAACGAGCTGCCGGGCACTGCCCCCGCTGCGCCTGCCCGCCTGCTGCTGGACCTCAGC  
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAAGTGGAGGCGCTGCGGCT  
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCAGAACCTCCACGACC  
 TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG  
 CGCTGCGGCTGGCCGGCAACACCCGCAT'TGCCAGCTGCGGCCCGAGGACCTGGCCGGCTGGC  
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC  
 TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC  
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA  
 CTTCGCCGCCAAGAAGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG  
 CCACCACCACCACAGCCACAGTGCCCAACACAGAGGCCCGTGGTGCGGGAGCCACAGCCTTGTCT  
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACGCGCCGCCACTGAGGCCCCAGCCGCCCTC  
 CACTGCCCCACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA  
 ATGGGGGCACATGCCACCTGGGGACACGGCACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG  
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG  
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC  
 AGCGTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG  
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCAACCA  
 GCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG  
 GCGAGGAGGCCTGCGGGGAGGCCCATACACCCCGAGCCGTCCACTCCAACCACGCCCCAGTCACC  
 CAGGCCCGGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC  
 GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGCGGGGGGGGCCATGGCAGCAGCGGCTCAGG  
 ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCTTGGAG  
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC  
 ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTACCCCTCCACGCAAAGCCCTACATCTAAGCCA  
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGTGCC  
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT  
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC  
 CTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC  
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGTGGGCTCTCCAC  
 TCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC  
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAACTGGAAAGGAAGATGC  
 TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTTATCT  
 GGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATGAA  
 GGCCTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

**FIGURE 16**

MCSRVPLLLPLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTPRDVPPDTVGLYVFENGIT  
 MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY  
 LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDANVE  
 ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL  
 AGLAALQELDVSNLSLQALPGDLSGLFPRLRLIAAARNPFNCVCPPLSWFGPWVRESHVTLASPEE  
 TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP  
 SPPSTAPPTVGPVPQPDQCPPSTCLNGGTCHLGRHHLACLCPEGFTGLYCESQMGQGTRPSPTP  
 VTPRPPRSLTTLGIEPVSPTSLSRVGLQRYLQGSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY  
 TVTQLRPNATYSVCMPLGPGRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV  
 LLAALAAVGAAYCVRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE  
 CEVPLMGFPGLQSPHAKPYI

**Important features:****Signal peptide:**

amino acids 1-23

**Transmembrane domain:**

amino acids 579-599

**EGF-like domain cysteine pattern signature.**

amino acids 430-442

**Leucine zipper pattern.**

amino acids 197-219, 269-291

**N-glycosylation sites.**

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

**Tyrosine kinase phosphorylation sites.**

amino acids 124-131, 337-345

**N-myristoylation sites.**
 amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,  
 594-600, 640-646





**FIGURE 18**

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSTTLTSDSVKDHTTAGRVVAGQIFLDSESEL  
 ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEPPKKVRKPALTAIEGTAHGEPCHFPFLFLDK  
 EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMQAEMMYQTGMKILNGSNKKSQKR  
 EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN  
 SSQAKALVYYTFGALGGNLIHMLVLSRL

**Important features:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 195-199, 217-221, 272-276

**Tyrosine kinase phosphorylation site.**

amino acids 220-228

**N-myristoylation sites.**

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

**Glycosaminoglycan attachment site.**

amino acids 267-271

**Microbodies C-terminal targeting signal.**

amino acids 299-303

**Type II fibronectin collagen-binding domain protein.**

amino acids 127-169

**Fructose-bisphosphate aldolase class-II protein.**

amino acids 101-119

**FIGURE 19**

AATTCAGATTTTAAGCCCATCTCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT  
 GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG  
 CCTCCTAATCCTCTGTGGT'TTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG  
 ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT  
 AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA  
 AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG  
 CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGTT  
 CCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA  
 CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG  
 TTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA  
 TATGCACTGGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC  
 ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC  
 TCGCCATTTGGGAGCAGCTGCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA  
 AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG  
 CATGGACCAGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA  
 TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA  
 GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA  
 AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT  
 CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT  
 CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT  
 GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA  
 TCTTTACCGTGGCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACCTTGTGAATGTTAAGT  
 ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAA

## **FIGURE 20**

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFBHIAACLTESG  
 STALKAETSERLRTVLLDVTDPENVKRTAQWVKVQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY  
 REPIEVNLFGLISVTNLMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMK  
 AFGVHVSCIEPGLFKTNLADPVKVKIEKKLAIWEQLSPDIKQQYGEYIEKSLDKLKGKNSYVNMD  
 LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-17

#### **Transmembrane domain:**

amino acids 136-152

#### **N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

#### **Glycosaminoglycan attachment site.**

amino acids 39-42

#### **N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

**FIGURE 21**

CTGAGGCGGCGGTAGCATGGAGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG  
 CACTCGCTTTCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA  
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA  
 CATTCAGAAATATATTCCATGCTATCAGCTTTTCTAGCTTTTATAATTCTTCAGGCGAAGTAAATG  
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCCGT  
 CGTCATTGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATTT  
 TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC  
 ATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTCACAGGGTACCTTTAGTGTTT  
 GCCAATCTGGGCTGTCTGAACAACCTGGGTATATAAACTGTATCAGGTTCTGTATGTCCACTGG  
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG  
 TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG  
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA  
 GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAAGACCCTCAGGAGA  
 ACATTTTTCTTTGTGAGGATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTGATGTGTT  
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACATAACCACCATCTCGATGTAGT  
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC  
 AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA  
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT  
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC  
 CTACATTTTGATCCTTTTAAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC  
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTACC  
 TGTTTGCAGTAATACACAGATAACTCTTAGTGATTTACTTCACAAAGTACTTTTTCAAACATCA  
 GATGCTTTTATTTCCAAACCTTTTTTTCACCTTCTACTAAGTTGTTGAGGGGAAGGCTTACACAG  
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT  
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT  
 GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA  
 GAAATTTATATGAAAATTTATCTGAGTCATTAAAATCTCCTTAAGTGATACTTTTTTAGAAGTA  
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT  
 AAAATTTAAAAA

## **FIGURE 22**

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI  
PCYQLFSFYNSSGEVNEQALKILSNVKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL  
VFLLLTPSIITESCSTHRLHSYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV  
QTHSSKFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA  
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDDVDNLT  
MVEHTDIPEASPSTPQIIKHKALDLDLDRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET  
DEEIEKMGFGGEYSRSPTF

**Important features:**

**Signal peptide:**

amino acids 1-19

**N-glycosylation sites.**

amino acids 75-79, 322-326

**N-myristoylation site.**

amino acids 184-154

**Growth factor and cytokines receptors family.**

amino acids 134-150

**FIGURE 23**

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA  
GCAGCGCGCAGCGAACGCCCCGCCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCCTGCCACCCCTTCCCTCCTTCCCC  
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCGCGGAAACCCGAGGTACACAGCCCGCGCCTCT  
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCGCGCCTGGCACCCGGGACCGTTGCCTGA  
CGCGAGGCCAGCTCTACTTTTCGCCCGCGCTCTCTCCGCCCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCTCCC  
TCCAGTCCACTCGCTAGTCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCGCTTCCCGTCCGGTCCCAAA  
GGTGGGAACGCGTCCGCCCGGCGCCGACCATGCGCACGGTTCGGCTTGCCCGCGCTTCTCTGCACCCCTGGCAGTGCTC  
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGCTTTACGTGTCCAAAGGCTTC  
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT  
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG  
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTTCTCAAAGAACTACTTGAATGCGAGAGAAATCCCTG  
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAAATTCGAGCTATTTAAAGATCTCTTCGTAGAGTTG  
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCTGGAGCGGATG  
TTCCGCTGGTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG  
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCGCTACTTTCGCTCAAGGC  
TTAGCGGTTGCGGGAGATGTCTGTAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG  
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC  
TGTTTGCCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGTTGGCAGAGAGGCTA  
GAGGGTCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT  
AATAGTGTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT  
TCCATCTCTGAAAGTGCCTTCAGTGTCTCGCTTCAGACCACATCACCCGAGGAACGCCAACCACAGCAGCTGGCACT  
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCGAGCAAC  
GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC  
CTGTTTGCACTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC  
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG  
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCCTGGGGCA  
CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG  
AAAAAGTGTTTCATCAAAAGTTAAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA  
TGACAACAATGTACAGTTTTTACTATGTGCCACTGGTTTAAAGAGTGCTGACTTTGTTTCTCATTAGTTTGGG  
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCAACCATGTAAACGTGGCTAACAGTGTAGGTACAGAA  
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTGTTTGTATGTTTTTCTCATTTCGTTTGTGGGT  
TTTTTTTCCAACGTGTGATCTCGCCTTGTCTTACAAGCAAAACAGGGTCCCTTCTTGGCAGCTAACATGTACGTATT  
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAGAAAAAGCCCAAAAGC

## **FIGURE 24**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST  
 CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH  
 LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE  
 QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL  
 VTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN  
 MQDNSVQVSQKVFQCGGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK  
 EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLEFAVTGNGLANQGNNPEVQVDT  
 KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG  
 KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

**Important features:**

**Signal peptide:**

amino acids 1-22

**ATP/GTP-binding site motif A (P-loop).**

amino acids 515-524

**N-glycosylation site.**

amino acids 514-518

**Glycosaminoglycan attachment sites.**

amino acids 494-498, 498-502

**N-myristoylation sites.**

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

**Glypicans proteins.**

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506



## FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC  
 CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT  
 TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA  
 TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC  
 CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCCTGAG  
 AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT  
 TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA  
 GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG  
 AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC  
 TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA  
 TGTTTTTCAAGATCATTTTGTGTTGTGCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT  
 GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT  
 AGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT  
 AAATGTCAAAAAAAAAAAAAAAAAA

## **FIGURE 26**

MKVLISLLLLLPLMLMSVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM  
TVSGLPKKQPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-myristoylation sites.**

amino acids 27-33, 46-52



## **FIGURE 28**

MDILVPLLQLLVLLLTPLPHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL  
 TGASGKVALLELGCGTGANFQFYPPGCRVTCIDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM  
 RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPTW  
 KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFSSKALICSFPSL  
 QLEQATHQPIYLPRLGT

**Important features:**

**Signal peptide:**

amino acids 1-23

**Leucine zipper pattern.**

amino acids 10-32

**N-myristoylation sites.**

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

## FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT  
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTTCTG  
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT  
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA  
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAA  
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAAGAAATT  
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC  
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

### **FIGURE 30**

MLLLTLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG  
EGEKVGDG

**Important features:**

**Signal peptide:**

amino acids 1-15

**Growth factor and cytokines receptors family:**

amino acids 3-18

**FIGURE 31**

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT  
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT  
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGAATCATGTCGG  
 GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC  
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT  
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA  
 CAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG  
 CTGCTCGTCTTGATTTTGTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC  
 AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA  
 TTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCAGGTT  
 ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTCGGGCATTCCGGTACATGTGGTTCGTACCATT  
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG  
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC  
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG  
 GATTCCGAGAATCATGTGCATGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT  
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC  
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC  
 ATTCAAAATCTTGTTCCAAGAACTCAAGTCACCTTACATCTATTAAGTCTTTGGAGACTTCATAA  
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC  
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC  
 CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACCTTTTCCTGTGTTTGTCTGTTGATC  
 TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA  
 AAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA  
 GGGAAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT  
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT  
 TTAAAAGACCTAATAAACCTATTCTTCCTCAAAA

## FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE  
LDTERENMKCVLGFIVSTGITAVLLVLI FVLKRRIKLTVELFQITNKAISSAPFLLFQPLWTFA  
ILIFFWWLVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA  
GAVVTCYFNRSKNPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG  
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD  
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA  
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

### Important features:

#### Signal peptide:

amino acids 1-20

#### Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

#### N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

#### N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348



TAACACAAACCTTCTTC

**FIGURE 34**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG  
 CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR  
 ESFIVLESKPKKGVTPSALTYSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA  
 TPTTLPRPSPAASTTIPRPQSVGHRSEQMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFAQP  
 VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI  
 GPAGPLMGVVQYGDNPATFNLKTHNTSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN  
 RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG  
 FYSLHVQSWFGLHKTLPVLRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN  
 LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL  
 FKKSHPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHDSFF  
 VDEFNLHQYVPRIIQNICTEFNSQPRN

**Important features:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 181-200

**N-glycosylation sites.**

amino acids 390-394, 520-524

**N-myristoylation sites.**

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,  
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

**Amidation site.**

amino acids 304-308

**FIGURE 35**

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC  
 CAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT  
 ACAGAAAACAACAAAAACTTAAGCTTTAATTTTCATCTGGAATCCACAGTTTTCTTAGCTCCCTGGACCC  
 GGTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCGAGTGTA  
 AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTA  
 GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGTCACCTCTGAGTTTCTTTGTGATGTGGTAC  
 CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAACGGATGTACTTCTATGAGTATGAGCCGATTTA  
 CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACCTGCTCTCATCAAAATCCATTCTGGTCATTCT  
 TGGTGACCTCCCAACCTTCAGATGTGAAAGCCAGGAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT  
 TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC  
 ATTTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA  
 ACCTGACCTTGAAGACATTATGGCATTCAAGGTGGGTAAGTGAAGTATCTTTAAACCTAAACCACTCAGA  
 GAAGTTTTTACAGGTTATCCTCTAATTGATAATTATCCTATAGAGGATTTACCAAAAAACCCATATTT  
 CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCCAGAGAT  
 TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAGAACCCATCAAGTTTGAAGATGTTTATGTCCGGAT  
 CTGTTTGAATTTATTAAGTGAACATTCATATTCCAGAAGACACAATCTTTTCTTCTATATAGAATCC  
 ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTCTTCCAAGGAGATCATCACTTTT  
 TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAACCTTCACATTCTACAAAAAGCCTAGAAGGACAG  
 GATACCTTGTGGAAGTGTAAATAAAGTAGGTACTGTGGAATTCATGGGGAGGTCAGTGTGCTGGCTT  
 ACACTGAACTGAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG  
 CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG  
 ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTACTGAGTTATAAGCTCA  
 CTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTTGTGTA  
 TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA  
 CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATT  
 ATTATTTAAATTTACTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG  
 TGAATCATCTTTTACATGCAACATTTCCAGTTACTTAACTGATCAGTTTATTATGATACATCACTCCA  
 TTAATGTAAAGTCATAGGTCATTATGTCATATCAGTAATCTCTTGGACTTTGTAAATATTTTACTGTGGT  
 AATATAGACAAGAATTAAAGCAAGAAAATCTGAAAA

## **FIGURE 36**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF  
 TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA  
 LSLEDEHLLYGDIIRQDFLDTYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFINTGNLVK YLL  
 NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV  
 KPIKFEDVYVGICLNLKVN IHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFS SKEIITFWQVMLR  
 NTTCHY

**Important features:**

**Type II transmembrane domain:**

amino acids 20-39

**N-glycosylation sites.**

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

**Glycosaminoglycan attachment site.**

amino acids 239-243

**Ly-6 / u-PAR domain proteins.**

amino acids 23-37

**N-myristoylation site.**

amino acids 271-277



**FIGURE 38**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT  
 IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGDCMRCGQ  
 VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII  
 KRVCGNERPAPIQSIGSSLHVLHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC  
 ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE  
 KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPHQLYSAAFQKQKLSAPTK  
 KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP  
 KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL  
 KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR  
 DLSTSFQESHITVAGWNVLADVRSPGFKNLTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA  
 SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLGLVSWSYDKTCSHRLSTAFTKVLPEKDWI  
 ERNMK

**Important features of the protein:****Signal peptide:**

amino acids 1-23

**EGF-like domain cysteine pattern signature.**

amino acids 260-272

**N-glycosylation sites.**

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

**N-myristoylation sites.**amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,  
474-480, 491-497, 638-644, 666-672**Amidation site.**

amino acids 56-60

**Serine proteases, trypsin family.**

amino acids 489-506

**CUB domain proteins profile.**

amino acids 150-167

## FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTATTAAACGTGGCTTAATC  
 TGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG  
 TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC  
 TGGTCTTGCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACCTCCTGTATTAGAACTCTGTA  
 AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAGATGGCTGTCCAGACGGCTGTGCGAGCC  
 TCACAGCCACGGCTCCCTCCCGAGAGGTTTCTGCAGCTGCCACCATTCTCTTAATGACAGACGAGCCTGGCCTAGACA  
 ACCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCACTGGACTCTGGCCGGAGCAACCGAACTA  
 GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA  
 GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGAAAACACCACTGCCCTGAAG  
 TCTTTCCAAGGTTGTACCACTGATTCCAGATGGTGAATTAACAGCATCAAGATCAATCGAGTAGATCCAGTGAAA  
 GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGG  
 TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC  
 ACAACTACGCTGTGCTCTCTCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAAGTCCGCA  
 GCAGGAACAATGGACAGGCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC  
 CCAGGAGCAGCTTGGAATAAACTGGTGCACAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGTGGATGGCG  
 GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA  
 GCCCAGAAAAGTGGGCTCATCTGATTGAGGCCAGTGAAGAGCGTGTTCACCTCGTGTCTCCGCCAGGTTCCGCGAGC  
 GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCAGGGCCAGGGGAGAGGAGCAACA  
 CTCCCAAGCCCTCCATCCTACAATTACTTGTCTAGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG  
 GCATGACCGTCGCAGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCACTAGTGTGAGCCCGAGGAG  
 TCATAAGCAGAGATGGAAGAATAAAAAAGGTGACATTTTGTGAATGTGGATGGGGTCGAACTGACAGAGGTGAGCC  
 GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC  
 CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCT  
 GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC  
 TGGGCTTCTGCATTGTAGGAGTTATGAAGAATACAATGGAACAAACCTTTTTCATCAAATCCATTGTGAAGGAA  
 CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATTTCTTCTGTGCTCAATGGTAGAGTACATCAGGAATGA  
 TACATGCTTGTCTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTCTTGGCCTGGCACTT  
 TTTTATAGAAATCAATGATGGGTGAGGAGAAACAGAAAAATCACAAATAGGCTAAGAGTTGAAACACTATATTTATC  
 TTGTGAGTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTGAGGAAAGTATGATCATCTAATGAAAGCCAGTT  
 ACACCTCAGAAAAATGATTCAAAAAAATTAAGTACTAGTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC  
 AACATTGTTTATATTTTCTATTCAATAAAAAGCCCTAAACAACATAAATGATTGATTGTATACCCCACTGAATT  
 CAAGCTGATTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT  
 AAAATATTTTAAATGCAATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAAGTTAA

**FIGURE 40**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRSQDGCPCDGCASLTATAPS  
 PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSEFKKINR  
 ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSSELSIRLV  
 GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIIILKVNMGDISNVPHNYAVRLLRQPCQVLWLTVM  
 REQKFSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDGGVAYRHG  
 QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVSRQVRQRSPDIFQEAGWNSNGSWSPG  
 PGERSENTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDR  
 IKTGDIILLNVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP  
 SDWSPSWMMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG  
 RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

**Important features:****Signal peptide:**

amino acids 1-15

**N-glycosylation sites.**

amino acids 108-112, 157-161, 289-293, 384-388

**Tyrosine kinase phosphorylation sites.**

amino acids 433-441, 492-500

**N-myristoylation sites.**

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,  
 467-473, 603-609



## FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT  
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT  
TTCCACCTTTCCTACAAATCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGTTGGGTGGGC  
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAGCAAAGGAGTTCATGGCTAATTTCC  
ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGT  
GACAAGTGTCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC  
TTTGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG  
CTTTACAGAGGGTCCGCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA  
CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG  
TAAAAAGTTTAAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG  
ACTGCTTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG  
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACCTGGGTACAGGTTACGTTACAGTGGAATTTTGG  
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAAGTACTGGGGAT  
GGGGAGGCCAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG  
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG  
GATGAAGCTCTTACACCAAGTGTCACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT  
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCAGAGTGGAATTTCTGGTTTGGTGCATGACCC  
TGGATCTTTTGGTGATGTTTGAAGAAGTATTTCTTTGTTGCAATAATTTTGGCCTAGAGACTTCAA  
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTCTTTTGTATTTCT  
TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT  
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA  
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT  
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA  
AGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAGGT  
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTTGGGAATGAAGA  
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC  
CTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT  
TTTTAAAGAGTTTTTGTAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT  
ATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

**FIGURE 42**

MGFNLT FHLSYKFRLLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN  
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH  
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFI FHDV  
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSYFGGVLTALSREQFFKVNGFSNNYWGWWGED  
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRDGLSSCSYKLV  
SVEHNPLYINITVDFWFGA

**Important features:****Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

# FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG  
 CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCCAACAG  
 ACGGGACAACCTGCAGAGCTGCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT  
 GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC  
 ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCCGTCCCCTCCC  
 TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTGTTT  
 TCCAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MALSSQIWAACLLLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI  
CIFCCGCCCHRSKCGMCCKT

**Important features:**

**Signal peptide:**

amino acids 1-24

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 58-59

**N-myristoylation site.**

amino acids 44-50

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 1-12

**FIGURE 45**

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC  
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT  
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG  
GACCTTCAACACAACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA  
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG  
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA  
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA  
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT  
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC  
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCCTGTCAGCAGAAACT  
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCTCCATG  
GTCTTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG  
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG  
AAACTCCTAACATATGCCCCCATCTTGAGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT  
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA  
TCTTAGACAGCAGTGCCTCCCCCTAAGTCTCTGCTCA

**FIGURE 46**

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP  
EGGTIIIVTQNRNRERVDFFDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK  
PKVTMGLQSNKNGTCVTNLTCCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC  
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE  
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLTMPDT  
PRLFAYENVI

**Important features:****Signal peptide:**

amino acids 1-22

**Transmembrane domain:**

amino acids 224-250

**Leucine zipper pattern.**

amino acids 229-251

**N-glycosylation sites.**

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,  
291-295

## FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG  
 ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT  
 TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT  
 ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG  
 CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGTCT  
 TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC  
 AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA  
 CTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG  
 ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG  
 CTTATCCACTTCTCAGTATTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCTGTTTGGGCT  
 CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG  
 TGTAGTTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

# THE BUREAU OF

Important features:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

amino acids 80-84, 132-136, 148-152, 163-167

amino acids 223-227

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

amino acids 207-218

amino acids 4-12



**FIGURE 49**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGGA  
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG  
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA  
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC  
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC  
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG  
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA  
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC  
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCTT  
TTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

## **FIGURE 50**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS  
SQKQHSPVPEKAIPILITPGSATTC

**Important features:**

**Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 36-59

**N-myristoylation sites.**

amino acids 41-47, 45-51, 84-90

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.**

amino acids 54-67

# **FIGURE 51**

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG  
 AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG  
 AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG  
 GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG  
 CGGAGAGGAAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA  
 GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGAGGGGCAGCTGGCTCTAAAGTCAGTGAG  
 GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC  
 AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA  
 TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG  
 GTGCCTGGCCACAGTGGTGCTTGGGAAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG  
 CCTTGGAGGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG  
 GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA  
 GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGAG  
 AGCCAGCAACCAGAATGAAGGGTCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA  
 ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC  
 AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG  
 TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT  
 CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA  
 CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATTCAGGG  
 CTTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG  
 GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT  
 GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACCTTGACACTTTCTGGAA  
 GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC  
 GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA  
 AACACCACCTCTCATCTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA  
 AA  
 AA

**FIGURE 52**

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS  
 EALGQGTREAVGTGVRQVPFGGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGV  
 GHSGAWETSGGHGIFGSQGGGLGGQCGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNF  
 GTNTQGAVAPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSGGS  
 SSGSSSGSSSGGSSGGSSGGSSGSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE  
 ARGSGESGIQGFRGQGVSSNMREISKEGNRLGSGSDNYRGQSSWGSGGGDAVGGVNTVNSETSPGM  
 FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

**Signal peptide:**

amino acids 1-21

**N-glycosylation site.**

amino acids 265-269

**Glycosaminoglycan attachment site.**

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

**Casein kinase II phosphorylation site.**

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

**N-myristoylation site.**

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,  
 90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,  
 159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,  
 236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,  
 253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,  
 283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,  
 298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,  
 325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,  
 387-393, 389-395, 395-401

**Cell attachment sequence.**

amino acids 301-304

## FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG  
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT  
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAAGTCCCGCCGGCTCCAGTGTTC  
CACAGCCCCCAAACGGAACGGTGTGGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGGC  
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT  
CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAGTGCCTCAGCTGCCA  
TTGCACCCAAGGATAATCTCTTCATCAGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG  
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT  
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG  
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA  
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT  
CTTGAGGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT  
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC  
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA  
AGCCAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGAGAGCCATGACACCAG  
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG  
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC  
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC  
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG  
CCTCATCGATATTATAGGGGTCCATCACAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC  
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTCCGCA  
GGGCCCAGGAACGTCATCGGGCAGGCGTTCCGCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT  
GCTGCTGCACTTCCGGTTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC  
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCACTGTTCTGAC  
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGTGTCAA

## FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFFQPPKRNWFWGHLG  
 LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDITRSITNASAAIAPKDNLFIRFLKP  
 WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHHLASEGSSRLDMFEHI  
 SLMTLDSLQKCI FSDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC  
 RLVHDETDVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF  
 MFGGHDTTASGLSWVLYNLAHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR  
 LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP  
 LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN  
 VGLQ

**Important features:**

**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

**Cytochrome P450 cysteine heme-iron ligand signature.**

amino acids 461-471

**N-glycosylation sites.**

amino acids 112-116, 168-172

**FIGURE 55**

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA  
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT  
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG  
TACAGCCTTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTTGCCGTGTGCTTGC  
ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCATATGGACAGAAGCTGGTGGACAGTTTT  
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTTGCTT  
GTGATTCGAACATTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA  
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAACTCTCCTCATGTACCTGTTTCCTC  
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA  
AAA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72



**FIGURE 57**

CGGCTCGAGCTCGAGCCGAATCGGGCTCGAGGGGCGAGTGGAGCACCAGCAGGCGGCCAATGCTCTGTCTGTGCCTG  
TACGTGCCGGTCACTCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC  
ATTTTCAAGCTCAGTGTCTTATCCCTCCAGGAATTTCCACCTACCGCCAGTGGGAAGCAGAAAATTGTACAAGCT  
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG  
CTGGTGTTTAAGATTTTGGACAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG  
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTTCTCAAGAGCATGGATAAAAACGGCCAGATGACCATCGACTGG  
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATCTCTACTGGAAGCATTTCCACG  
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA  
CACCTGGTGGCAGGAGGTGGGGCAGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG  
AGGTCATGCTCCCGCAGCAACACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG  
TCACTCTGGCGGGCAATGGCATCAACGTCCTCAAAATTGCCCGGAATCAGCCATCAAAATTCATGGCTATGAGCAG  
ATCAAGCGCCTTGTGGTAGTGACCAAGGAGACTCTGAGGATTCAGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCC  
ATCGCCCGAGGACAGCATCTACCAATGGAGGTCTGAAGACCCGGATGGCGCTGGGAAGACAGGCCAGTACTCAGGA  
ATGCTGAGTGGCCAGGAGGATCCTGGCCAGAGAGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTG  
GGCATCATCCCTATGCCGGCATCGACCTTGCACTCTACGAGACGCTCAAGAAATGCTGGCTGCAGCACTATGCAGTG  
AACAGCGCGGACCCCGCGTGTGGTCTCTGGCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC  
CCCTGGCCCTAGTCAGGACCCGGATGCAAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCTC  
TTCAACATATCTTGCAGACCGAGGGGGCTTGGGGTGTACAGGGGGCTGGCCCCCACTTCATGAAGGTGATCCCA  
GCTGTGAGCATCAGCTACGTGGTCTACGAGAACCCTGAAGATCACCTGGCGTGCAGTCCGGGTGACGGGGGAGGGC  
CGCCCGCAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT  
AAGCTGTCTCGAGCCAAGCTGTGAAAACCTTAGACGCACCCGCAAGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT  
GTCTGTGACCCAGCAGACCCTCCTGTTGGTTCCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAAGG  
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGAGGCTGGAGGCCGGCT  
TAGTTCTTCCATTTACCCCTTGACGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC  
CCTCTTGTGCTGCTGCTGTCTGCTGAGGTAAGGTGGGAGGAGGGGTACAGCCACATCCACCCCTCGTCCAATCCC  
ATAATCCATGATGAAAGGTGAGGTACGTTGGCCTCCAGGCCCTGACTTCCCAACCTACAGCATTGACGCCAATTTGGC  
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG  
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT  
GGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA  
CCATGTTGAGGGCGAAGGGCAGAGCGTTTGTGTGTTCTGGGAGGGAAGGAAAGGTGTTGGAGGCCCTTAATTATGG  
ACTGTTGGGAAAAGGGTTTGTCCAGAAGGACAAAGCCGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG  
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTCTCTGTCCAACCCAGCAGGGGCGCAGC  
GGGACCAGCCCCACATTCACATTGTGTCACTGCTTGAACCTATTATTATTGATTATTGAAACAGAGTTATGTCTCT  
AACTATTTTATAGATTTGTTTAATTAATAGCTTGTCAATTTCAAGTTTCAATTTTATTCATATTTATGTTTATGTT  
GATTGTACCTTCCCAAGCCCCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCTTGGGCCGTGCACTGACATCT  
GTCCAGAGAAATTCCTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG  
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTTACTGGGGCGTGGAGAGAGAGGGAGAACCTCAAT  
AACCTTGAAGGTGAATCCAGTTATTTCTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAG  
TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGCTCCCTCCAGC  
CTTCTGCTGCCCTTGTCTAACAAATGCCGGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAATGACCTGA  
TGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATTTGTTATATATGAACATATAACTGGAGTCGTCAAAAAG  
CAAAATTAAGAAAGAAATGGACGTTAGAAGTTGTCATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAGCTGAAAAA  
AA

## **FIGURE 58**

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG  
QLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQAEKILKSMDKNG  
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRLVAGGG  
AGAVSRTCCTAPLDRKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK  
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIIYPMEVLKTRMALRKTQYSGMLDCARR  
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC  
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV  
VYENLKITLGVQSR

**Important features:**

**Signal peptide:**

amino acids 1-16

**Putative transmembrane domains:**

amino acids 284-304, 339-360, 376-394

**Mitochondrial energy transfer proteins signature.**

amino acids 206-215, 300-309

**N-glycosylation sites.**

amino acids 129-133, 169-173

**Elongation Factor-hand calcium-binding protein.**

amino acids 54-73, 85-104, 121-140

## FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC  
 TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG  
 CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT  
 GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTGAACCTGACATCAAACCTTCTGATAT  
 CGTGATAACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCCAAAGATG  
 AGCTGTCTGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTCGTGATCAAGTGATAGTT  
 GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT  
 CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG  
 AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC  
 CAGCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC  
 CAGCTTTGAGCTGAACCTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA  
 TCAACAACACATACTCCTGTATGATTGAAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG  
 ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT  
 CTCTTCTTTCTTTGCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT  
 GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC  
 CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG  
 AGCAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT  
 CTATCTTCAAAGACATATTAGAAGTTGGGAAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA  
 GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT  
 CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTTATATGTGC  
 TGTAATGTTGCTCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC  
 AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG  
 GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCAAAGGTGCCTTGGCTTC  
 TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT  
 CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA



**FIGURE 61**

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG  
 CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT  
 GGGCTACCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACCAGCTGG  
 TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAATCCAGGAGGACCATATG  
 GCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGGCCAGCAGCCT  
 GGGCTTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC  
 GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA  
 ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA  
 GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAATTCATCCAGCAGTGGAAGAACCTCTT  
 CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC  
 AAATGGGCTACAACCTGAGCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT  
 GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGTGACAGA  
 GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA  
 CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT  
 TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC  
 ATTCTCCCTTGCTTGATGCAAACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA  
 AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA  
 ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCCAGGAGCAGGTCTTGTAAATGG  
 AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC  
 GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCAAAGTGGAAT  
 CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT  
 CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT  
 CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCT  
 TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTTC  
 ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

**FIGURE 62**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG  
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH  
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR  
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK  
DTAVQGNIRLSFEDEFVTMTASRML

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-  
66, 70-75, 78-83, 83-88, 87-92, 110-115

**FIGURE 63**

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC  
 TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC  
 AACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC  
 TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG  
 GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGGCTGCCTGCCATG  
 GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA  
 AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGCCCTT  
 TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC  
 TCGGAGTCCAGACGACTGCCCCGTCTAATTCACTGGGAGCCGGGGAAAAATCCTTTCCCAACG  
 CCCCTCCCTGGTCTCTCATCCACAGGGTTCGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG  
 TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA  
 ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG  
 CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG  
 GCAGCTGGGGGAATATTCATCTATAACCCAGGTATCAATAACCCATTTCTCTCTGGAGTTCTCCGC  
 CCTCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG  
 GGGCTAGAGCACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG  
 TGTGGGCTCAATCCAGGCCCTGTTAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCCTCCCC  
 TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
 AA

## **FIGURE 64**

MQGRVAGSCAPLGLLLVLHPLGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGSPNSEHPQPAL  
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVQWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA  
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP  
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWNINRYPGGS  
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

**Important features of the protein:**

**Signal peptide:**

amino acids 1-26

**Casein kinase II phosphorylation sites.**

amino acids 56-59, 155-158

**N-myristoylation sites.**

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,  
259-264, 269-274, 270-275, 280-285, 281-286, 305-310



## FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC  
 TGGGCTGCCCCCTTGTCTCCTCTTGACCCCTCCTGGCAGCTCACATGGAACAGGGCCGGGTATGA  
 CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCTTGAA  
 TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC  
 AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGCCCC  
 GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCTC  
 CTGAGTGGCAATAAATAAAATTCGGTATGCTG

**FIGURE 66**

MGSGPLVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTL  
HHARSQHHVVCNT

**Important features:****Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 37-41

**N-myristoylation sites.**

amino acids 15-21, 19-25, 60-66

**FIGURE 67**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGC  
CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC  
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCGGGCTGGGGCTGCTTCTGGCG  
CTGGGCCTGCCGTTCTGTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC  
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG  
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG  
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA  
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT  
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA  
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA  
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACCTTTATATATTTATATAAAATTAG  
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

**FIGURE 68**

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVVFS  
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

**Important features:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 56-80

**N-glycosylation site.**

amino acids 36-40

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 86-90

**Tyrosine kinase phosphorylation site.**

amino acids 86-94

**N-myristoylation sites.**

amino acids 7-13, 26-32

**FIGURE 69**

GCCAGGAATAACTAGAGAGGAACA**ATG**GGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCACCAG  
 TCAAAATCTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA  
 GATGAAAAATAAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA  
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAATGAA  
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTTTACA  
 GAATGTGGAGAGAAAGGCCAATACATTCACTTCAACCTGACCTTCTACTTGGAAAAAACAAAATGAATATGGACCA  
 CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC  
 TACCGTGTCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT  
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATGGAAAAAGATTGTCAATTCTTT  
 CCTGATAAAGTACAACAGAAAAAGCATCCATAATGTTATGCAAAAGTATTGATTCTGTTGTAATTTTGTAAACGAA  
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT  
 TCTGAGGATTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA  
 AGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGTAAGGACCGCCTAAATCGAATGAATCAAGCAGCA  
 AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA  
 AATAAGCTAATCCAAATAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA  
 ACTTCCATCTGCTCTGGAATTAATATGCATTTCAAGTGATTGGAGAGCTACATTCCCAACTCGATGGATCCGAAGTA  
 CTGCTGCTGACTGATGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCAT  
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCAATTTTATGTT  
 TCAGATGAAGCTCAGAACAAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAA  
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGAATAGTAATGCCTGGATGAACGACACTGTCATAATTGATAGTACA  
 GTGGGAAGGACAGCTTCTTCTCATCACATGGAACAGTCTGCCTCCAGTATTTCTCTCTGGGATCCCAAGTGAACA  
 ATAATGGAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAACTGCAAAAGTGGGCACT  
 TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAACATTAACATTACAGTAACCTCTCGAGCAGCAAAATCTTCT  
 GTGCCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCAGCCCAATGATTGTTTACGCAGAA  
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCACAGAATGGACATACAGAAGTT  
 TTGGAACCTTTGGATAATGGTGCAGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT  
 ACAGAAAATGGCAGATATAGCTTAAAGTTCGGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGGCCTCCA  
 CTGAATAGAGCCCGGTACATACCAGGCTGGGTAGTGAACGGGGAAATGAAGCAAAACCCGCAAGACCTGAAATTTGAT  
 GAGGATACTCAGACCACCTTGGAGGATTTCAAGCCGAACAGCATCCGGAGGTGCATTGTTGGTATCACAGTCCCAAGC  
 CTTCCCTTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCAGAGTTCATGAGGATAAGATTATTCTT  
 ACATGGACAGCACCAGGAGATAATTTGATGTTGGAAGGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT  
 GATCTAAGAGACAGTTTGTGATGCTCTTCAAGTAAATACTACTGATCTGTCAACAAAGGAGGCCAACTCCAAGGAA  
 AGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTTATTGCCATTTAAAGTATAGATAAA  
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTGTATCCCTCAAGCAAATCCTGATGACATTGAT  
 CCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAAGTCATAATCTGGAGTTAATTTCTACGCTGGTATTG  
 TCTGTGATTGGGTCTGTTGTAATTTAACTTTATTTTAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC  
 AAGTAGACCTAGAAGAGAGTTTAAAAAACAAAACAAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT  
 GTGTGATCATAACTCATAAAAATAATTTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAAACTCATGGATA  
 TGTAAAACTGTCAAGATTAAAAATTAATAGTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAACAAAG  
 ATCCTTTTTCATACTGATACCTGGTGTATATTATTGATGCAACAGTTTCTGAAATGATATTTCAAATTCATCAA  
 GAAATTAATAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAAATAACAACATTTGGAAAAA  
 AA

**FIGURE 70**

MGLFRGFVFLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIEQIEDMVTTASTYLFE  
 ATEKRRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY  
 IHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN  
 RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMSIDSVVEFCNEKTHNQEAP  
 SLQNIKCNFRSTWEVISNSEDfKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRLNR  
 MNQAAKHFLQLTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK  
 YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG  
 SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL  
 ITWNSLPSSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR  
 AANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA  
 GADSFKNLGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP  
 RPEIDEDTQTTLDEFsRTASGGAFVVSQVPSLPLPDQYPPSQITDLATVHEDKIIILTWAPGDN  
 FDVGKVQRYIIIRISASILDLRDSFDALQVNTDLSPKANSKESFAFKPENISEENATHIFIAI  
 KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI  
 VNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,  
 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

**FIGURE 71**

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGGTGA  
CAACAGGTGTCATCTTTTGTATCTCGTGTGTGGCTGCCTTCTATTTCAGGAAAGACGCCAAGGTAATTTTGACCCA  
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA  
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTTGTGGTTGGAGGAGAGAACCCTTTGTGGGGCTGCCTTCTTAGCA  
GTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAAAGAAAGGAAAGGTCCTTCTGCTGTTGGCTGCACATCAGGAA  
GGCTGTGATGGGAATGAAGGTGAAACCTGGAGATTTCACTTCAGTCATGCTTCTGCCTGCAAGATCATCCTTTAAA  
AGTAGAGAAGCTGCTCTGTGTGGTGGTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC  
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCAGGGAAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC  
ACCGGTTCTGGACGCATGGCTGATTCTGCAATGATGATGCTTCCGGGGGGCTGCTTGCCTGGATTTCCCGGGTGGTG  
GTTTTGCTGGTCTCCTCTGCTGTGCTATCTCTCTGTACATGTTGGCTGCACCCAAAAGGTGACGAGGAGCAG  
CTGGCACTGCCAGGGCCACAGCCACCGGGGAAGGAGGGGTACCAGGCCGCTCCTCAGGAGTGGGAGGAGCAGCAC  
CGCAACTACGTGAGCAGCTGAAGCGGCAGATCGCAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGG  
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC  
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG  
CCTTTTCGATAGCTTACTCTACAGAAAGGTGACAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAGCCTGTG  
AGGAAGGACAAAGCGGATGAGTTGGTGGAAAGCCATTGAATCAGCCTTGGAGACCTGAACAATCCTGCAGAGAACAGC  
CCCAATCACCGTCTTACAGGCCCTCTGATTTCTAGAAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT  
GAGCTCACCTTCAAAGGGGACCAAAACAGCAATTCAAACGGCTCATCTTATTTCGACCAATTCAGCCCCATCATGAAA  
GTGAAAAATGAAAAGCTCAACATGGCCAAACAGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC  
CGGCAGTTTATGCAGAAATTCAGGGAGATGTGCAATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG  
AAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTCAGGAACTTACCTTCATC  
CAGCTGAATGGAGAATTTCTCGGGGAAAGGGACTTGATGTTGGAGCCGCTTCTGGAAGGGAAGCAACGCTCCTTCTC  
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAG  
GTATTTTATCCAGTTCTTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGAA  
CAGCAGCTGGTCAATAAAGAAAGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGCTCAGTATCGGTGAGAC  
TTCATCAATATAGGTGGGTTTGTCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC  
CAGAGCAACCTCATAGTGGTACGGACGCCTGTGGCAGGACTCTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG  
CTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG  
GTGTTAGGACAGGATAGAGGCTCACCTTCGCAAAACAGAAACAGAGACAAGTAGCAAAAAACATGAACCTCCAGAG  
GAAGGATTGTGGGAGACACTTTTCTTCTCTTTTGCATTTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA  
GGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTTGGGCTTTTTTACAACAGA  
AATCAAAATCTCCGCTTTCCTGCAAAAGTAACCCAGTTGCAACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG  
AGATTATAAGCCTAATGGTGTGGAGGTTTGTGAGTGTGTTTACAATACACTGAGACCTGTTGTTTGTGCTCATTGA  
AATATTCATGATTTAAGAGCAGTTTGTGAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAATATCAGAAGGCAGGAGAGGATAGGCTTATTATGATACT  
AGTGAGTACATTAAGTAAAAATAAATGGACCAGAAAAGAAAGAACCATAAATATCGTGTCAATTTTCCCCAAGAT  
TAACCAAAAAATAATCTGCTTATCTTTTGGTTGCTTTTAACTGCTCCGTTTTTTCTTTTATTTAAAAATGCACT  
TTTTTCCCTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC  
ATTTTTATATTTTTTAAAGAGATACTTTGAGATGCATTATGAGAATTTTCAGTTCAAAGCATCAAATTTGATGCCATAT  
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCACTGAGACATAGGGAAGGAATGGTTTGTACT  
AATACAGAGCTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC  
ACTTTCTGCTTTACAGAAAAGGAAACTATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACACATTTT  
CTCCTCAGAGTAGGGACCGCTTCTTACCTGTTAAATAAACCAAGTATACCGTGTGAACCAACATCTCTTTTC  
AAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATTGT  
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG  
TAACCTGAATATTTTTTAAATTAAGCAGTTCTACTCAATCACCAAGATGCTTCTGAAAATTTGATTTTATACCATTT  
CAAATATTTTTTAAAAATAAATACAGTTAACAATAGAGTGGTTTCTTCAATCATGTGAAAATTTATAGCCAGCACAG  
ATGCATGAGCTAATATCTCTTTGAGTCCTTGTCTGTTTGTCTCACAGTAACTCATTGTTTAAAGCTTCAAGAAC  
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG  
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATGATTGTGGATATGAA

**FIGURE 72**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW  
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK  
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNNPA  
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNEKLNMAN  
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF  
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY  
NPGIIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDLDIKGWGGEDVHLYR  
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPQYKMCMQSKAMNEASHGQLGMLVFRHEIEAHL  
RKQKQKTSSKKT

**Important features:**

**Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 315-319, 324-328

**N-myristoylation sites.**

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

**Amidation site.**

amino acids 377-381



[illegible]

## FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNASPIDIQTDSVTFDPLPALQ  
PHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGSEHQINSEAT  
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP  
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV  
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS  
VVFTSAQATTEA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69

**FIGURE 75**

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC  
 TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCCTACA  
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA  
 ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGCTCCTGACAGCTCC  
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGA  
 AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC  
 ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC  
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT  
 CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTT  
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT  
 GATTTTGATTTATGGAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA  
 ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA  
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA  
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAATTTTTTGTGACTCTGAAG  
 AAAACACGGAAGGTACTTCTCTCACCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAA  
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCTGAAGAGCAGGA  
 GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG  
 TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCCTGGCG  
 CAGGAGCACACAGACTCGGAGGAGGGGCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA  
 TCCCCAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG  
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT  
 CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA  
 TGTGCAGATGGAAAACTGATGCCAAACACTTCCTTTTGCTTTTGTTCCTGTGCAAACAAGTGAG  
 TCACCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTCACTGT  
 CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA  
 GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTCTATGCAGAGAA  
 AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAAGTGGGTGT

**FIGURE 76**

MSYNGLHQRVFKELKLLTLCSISSQIGPPEVALTTDEKSI SVVLTAPEKWKRNPEDLPVSMQQIY  
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK  
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEFDKRFFVPAEK  
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC  
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLESQA  
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS  
EGCEPSEGDLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

**Important features:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 140-163

**N-glycosylation sites.**

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

**FIGURE 77**

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC  
 CTGGGAAGATGCCGCGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC  
 CAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC  
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC  
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC  
 ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA  
 GGAGTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA  
 TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC  
 CCCACCCGCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA  
 TAAGTCTCCTTCCTGGTGAACGCCTTAGCTAAGCAGGTCATGAACCTCCTAGTGCCATCCCTGC  
 CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC  
 CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA  
 TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA  
 AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCTGACAATGCCACCCTGGACAACATCCCG  
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
 ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC  
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC  
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT  
 TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT  
 TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG  
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA  
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG  
 CCTTGGGATTCGAGGCAGCTGAGTCTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC  
 TTGTGGAAACCCAGCTCTCCTGTCTCCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG  
 GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACTTG  
 CCTGTGAAAAA

**FIGURE 78**

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQPLLSAMREK  
PAGGIPVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE  
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL  
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT  
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHRKSSIGLIN  
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT  
KGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG  
FEAAESSLTGDALVLTASLWKPSSPVSQ

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

**Glycosaminoglycan attachment site.**

amino acids 412-415

**LBP / BPI / CETP family proteins.**

amino acids 407-457

## FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC  
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT  
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTGGGCACACTGGTTGCCAT  
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT  
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC  
ACCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT  
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC  
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCCTGGGATTC  
ATTCTGTGTGCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG  
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG  
CTGGAATCATCTCTGCTTTTCCTGCTCATCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC  
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA  
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG  
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT  
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG  
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC  
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA  
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA  
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT  
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC  
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA  
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG  
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

**FIGURE 8o**

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD  
IYSTLLGLPADIQAAQAMMVTSSAISLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL  
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY  
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

**Important features of the protein:****Signal peptide:**

amino acids 1-24

**Transmembrane domains:**

amino acids 82-102, 117-140, 163-182

**N-glycosylation site.**

amino acids 190-193

**PMP-22 / EMP / MP20 family proteins.**

amino acids 46-59



# **FIGURE 81**

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT  
 CCCGCGTTCTCTTTCCACCTTTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCTT  
 GCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG  
 GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG  
 ACCAGCGGCCTGACCCTGGGGAAGGATGTTCCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA  
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT  
 CCATGGGAAGAGATACTCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT  
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCT  
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC  
 TCCCTCTGGACTCCGGGCCCCACCAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG  
 AGATCTTCAGTGCCCATGAGCTGTTCCTTCCCGCTGCCCAACAGTGTGTCTCTGCAGCTGC  
 ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACAGGCTGCCAGCACCCCTCCC  
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA  
 GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG  
 AGAGGGCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT  
 CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT  
 GTGTGCATGGCGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCCTTCGGC  
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC  
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG  
 ACAAGCAGACCCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC  
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC  
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG  
 GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG  
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCCT  
 GGAACGTCTTCTAGCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG  
 ACATAACAAAGACC**TAA**CAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA  
 TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 82**

MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE  
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF  
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR  
HPQDPCSSDAGRKRGPOTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTYS  
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPCRHPEKVAGKCKICPEDKADPGHSE  
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH  
SQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPGAEGHGQSRQSDQDITKT

**Signal peptide:**

amino acids 1-25

**FIGURE 83**

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGCAGAGCCTCTCC  
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA  
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC  
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCCTGGTGGGGGAG  
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG  
 CCAGTTCTTAGCGTGGTCCACCTCTACAGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC  
 AAGGCAGGACAAACTGGTGAAGGATTCTATGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT  
 GTGTTGGATGCTGGCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT  
 ACAGGTGTCAGCACTGGGCTCAGTTCCCTCTCATTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC  
 TCTGTCACTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTTGTCC  
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA  
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG  
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA  
 TTTTTTGGCATTGTGGAAGATTTTCTCTCCAAATTCAGTGGAAAATCCAGGCGGAAGTGGAGTGTG  
 GAGAAGAAAGCACGACAGGCAGAAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG  
 AGACGGCTCACCCGAAGCTCTGCGTTCTGATCTGAAAAGTGAACCCATAGAAAAGTCCCCAGGAGGTG  
 CCTCACTCTGAGAAGAGATTACAAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA  
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA  
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT  
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA  
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG  
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAAGTCCCATAGTCATCTGC  
 CCAGTCAACCAGGAATCAGAGAAAGAGGCCCTTGTGGCAAAGGCCCTCTGCAATCCAGAGACAAGCAACAG  
 TGAGTCCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT  
 TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGCCAAGGTGGCTTCCA  
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG  
 CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG  
 AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG  
 ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCATTATATTACACTTTCAGTAAAAAA

## FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH  
 LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ  
 VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL  
 TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK  
 FQWKIQAELDWRKKGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTTHRKAPQEVPHSEKRF  
 TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT  
 LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT  
 PIVICPVTQESEKEASWQRASAIPESTNSSESSQATTPLPRGEM

**Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 239-255

## FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT  
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA  
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT  
TACCCTGGCCCAGTAGTTCATGGCTACTGGTTCGGGAAGGGGCAATACAGACCAGGATGCTCC  
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG  
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA  
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA  
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC  
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCCTATGATCTCCTGGATA  
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCCCTCATCCCACA  
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA  
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCCTCAGAACTTGACCATGACTGTCTTCCAAGGA  
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT  
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA  
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT  
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG  
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA  
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTC  
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG  
CTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGTCCAGTATGCATCCCTCAGCTTCCAGATGGTG  
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG  
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA  
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG  
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC  
TCCCTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

**FIGURE 86**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYPGPVVHGYWFREG  
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY  
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTTPMISWIGTSVSPLDPSTTRS  
SVLTLPQPQDHGTS LTCQVTFPGASVTTNKT VHLNVSYPQNLMTV FQGDGT VSTVLGNGSSL  
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP  
LGSQQVYLNVS LQSKATSGVTQGVVGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE  
DANAVRG SASQG PLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE  
YSEIKIHR

**Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 351-370

**FIGURE 87**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC  
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC  
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT  
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA  
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG  
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG  
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC  
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT  
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG  
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT  
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC  
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTACCCCTAT  
GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA  
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG  
GATACTTTCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG  
TGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA  
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA  
AAAAAA

## FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI  
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT  
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG  
EGKCWTDNGPVIPIVVYDFGDAQKTASYISPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN  
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

**Important features:**

**Signal peptide:**

amino acids 1-16

**N-glycosylation site.**

amino acids 163-167

**Glycosaminoglycan attachment sites.**

amino acids 74-78, 289-293

**N-myristoylation sites.**

amino acids 76-82, 115-121, 124-130, 253-259, 292-298





## **FIGURE 90**

MGRVSGLVPSRFLTLLAHLVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA  
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV  
TVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

**FIGURE 91**

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC  
 TCGCTGCTGCTTCGTGTTCCCTGGTGACGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG  
 GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCCT  
 CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT  
 GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC  
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG  
 GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT  
 CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG  
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG  
 GAGTGGGAGAAGGTAGAACGGGGCCCGCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG  
 CTCCCAGACACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG  
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCCAGATTACAACATAC  
 CATAGTGATACCCCTACTACCCATCTGGGTGACCCCGGGCAGGCCACAGAGGCCAGGCCAGGGC  
 TGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGCAGGGAAGGGGTGGGCCCTC  
 AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
 AGAAAGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG  
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG  
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC  
 CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA  
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAGCAGGGTGGGGTGGG  
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCGTGTGCTGAGCATGG  
 CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC  
 CAGGCCACCCCTTTCCAAAATCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG  
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC  
 CGCGTGTCTGTGTCTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG  
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTTCTCCCGGATCTGGATGGCGC  
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC  
 TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT  
 TGTGTGAAGAATCGTGTCTTGGAGCAGGAATAAAGCTTGCCCCGGGGCA



# **FIGURE 93**

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC  
 CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT  
 CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA  
 AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA  
 ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC  
 AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA  
 GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT  
 GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT  
 ATTCTGGGGCATTGTATTTTTTGTATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC  
 TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG  
 TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG  
 AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT  
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACCTGTGCCT  
 TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

## **FIGURE 94**

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG  
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV  
FSFVNTLSDSLGPSTVGIHGDSPOFFLYSAFMTLVIILHVFVGIVFFDGCEKKKWGILLIVLLT  
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKCLLCQDKNFFLYNQSR

**Important features of the protein:**

**Signal peptide:**

amino acids 1-19

**Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

**Glycosaminoglycan attachment site.**

amino acids 120-123

**Sodium:neurotransmitter symporter family protein**

amino acids 31-65

**FIGURE 95**

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG  
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC  
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG  
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTAAGTCTGATACCATTAACA  
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC  
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC  
CAATTTTGTACACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC  
TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC  
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG  
GAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT  
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA  
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC  
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT  
TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA  
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT  
TAAACATATTTGGAACCTGGAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIGURE 96**

MRSTILLFCLLGSTRSLPQLKPAIGLPPTKLAPDQGTLPNQQSNQVFPSLSLIPLTQM  
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGGLNVQQQLHHPVLPFIQVTQLGAQGTILSSEE  
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16





## FIGURE 98

MVPAWLWLLCVSVPQALPKAQPaelSVeVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT  
 EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI  
 YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG  
 STSLDHALERTYQLLVQVKMDGDAQSGHQATATVEVSIIESTWVSLEPIHLAENLKVLYPHHMAQ  
 VHWSSGDVHYHLESHPPGPFfeVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL  
 VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA  
 FQVDPTSGSVTLGVLPRLAGQNI LLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG  
 PISLPEDVEPGTLVAMLTAIDADLEPAFRMLDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY  
 EAAPSHEVVVVVQSVAKLVGPGPGPGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT  
 IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTAPVPS  
 QYLCTPRQDHGLIVSGPSKDPDIASGHGPYSFTLGNPTVQRDWRLQTLNGSHAYLTALHWWEP  
 REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGIILVGTTLVAIGIFLI  
 LIFTHWTMSRKKDPDQPADSVPLKATV

### Signal peptide:

amino acids 1-18

### Transmembrane domain:

amino acids 762-784

## FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG  
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA  
 TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC  
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG  
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC  
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG  
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA  
 GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG  
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG  
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCCACCAACT  
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC  
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGTGTCC  
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA  
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT  
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA  
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG  
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCCAGCCACCAACT  
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACACCTCC  
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC  
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA  
 CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA  
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG  
 CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATCTGAGTCCAGCACAACTCCAGTGGGGCCA  
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA  
 ATGCACACAACCTCCCATAGTGCATCTACTGCAGTGAAGGCAAAGCCTGGTGGGTCCCTGGT  
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC  
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT  
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG  
 GTGGAGTCCTAACCTGGTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA  
 ACAGCGGGCCCCTGAAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA  
 CCCAAGACCTGGTTTCCTTTTCATTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAA  
 AATCTTGAAGAAGGTATTCCTCACCTTTCTTGGCTTTACCAGACACTGGAAGAGAATACTATAT  
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG  
 CCCCAGGGTGGGTATCTAGCTCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC  
 CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 100**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVSSGASTATNSGSSVTSSGVSTATISGS  
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS  
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA  
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES  
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS  
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA  
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES  
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVA  
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVS  
AMEMSGRNSGP

### **Signal peptide:**

amino acids 1-20

### **Transmembrane domain:**

amino acids 510-532

## FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA  
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG  
GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT  
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGTCTCTGAACCCATGGTCAATT  
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT  
GGTGTTCATGGAAGGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGGGAAGCATTT  
TCATGCTGAGTCCC'TTTTACCTTTGATGTTTGTAAACCCATCTTGGTATCGCTGGATCAACAAC  
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGT  
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA  
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAAATT  
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTTGGATTGTTGGGCCATGCAGGCTGCTGCCTATAT  
CTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCAAGACATGATTGATTACTTTTGTG  
ATATTCACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACGAAAAACAGCAAG  
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG  
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG  
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT  
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA  
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG  
GGGAGAAGAATTTTATTTTACCGGACAGAGTGTCAATTCACCTTGCAAGTCTGAACTCAGGGTC  
CTTGTGGTCAAATTGCTCTCTTACTGTATTGGACCCTGTTTACGCCCTGCAATGTGCCTACTCAT  
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA  
GAATATTTGGTGGACTGGAGATCATAGAATTGCATGTTACCGACTTTTACACAAACAGCCACAT  
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG  
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTCGAATATTTCTTACT  
GCCATCATTATTTGTTAAAGATATTTTGCACTTAATTTTGTGGGAAAAATATTGCTACAATTTT  
TTTAATCTCTGAATGTAATTTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT  
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

## **FIGURE 102**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIIFMLSP  
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM  
FLWNCMLRYSYLRLKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP  
LQLLIFFEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA  
YPHNIPQSEKHLQGDFFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF  
YFTGQSVIPPCKSELRLVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG  
LEIIEELACYRLLHKQPHLNSKKNE

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 44-63, 90-108, 354-377

**FIGURE 103**

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC  
 TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACAGGAATA  
 TCC**A**TGGCTTTTGTGCTCATTITGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT  
 CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT  
 TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTCTTTCAGGAATCAGTTCCATGCTGTGGTC  
 CACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA  
 GTTTGTGAAGGACTCCATTGCAGGGGGGGCGTGTCTCTCTAAGGCTAAAAACATCACTCCCTCGG  
 ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG  
 CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCATCGTGGGATATGTTGACGGAGGTATCCA  
 GTTACTCTGCCTGTCTCAGGCTGGTTCCTCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC  
 AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC  
 ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCATCCACCTTGTCTGAGCAGAGTCAATGA  
 GGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCTGGCTTCTA  
 TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC  
 AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG  
 AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG  
 TTTCTGATCTGAAAACGTGAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA  
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA  
 CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA  
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT  
 ACATTCAATCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTCT  
 GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTTGC  
 TGACATGTGAGTTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAAG  
 GGGACTCCCATATTCATATGTCCAGTGTCTGGGGAT**G**AGACAGAGAAGACCTGCTTAAAGGGC  
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCCGACAGGTGGCCCCAGCTTCTCT  
 CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC  
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG  
 TCAGAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTAGTTTGTGAAAA  
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTACCGG  
 ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGAGGTT  
 TGAGGGCACAGTGTGTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT  
 TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC  
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCAGCAT  
 TACCTGATACCAAACAGGCAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC  
 CTCATTAACACAGACACAAAAATTTCTAAATAAAATTTTAAACAAATTAACATAAACAATATATTTA  
 AAGATGATATATAACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTAATATTTAAAT  
 ATCAACCAGTGTAAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

**FIGURE 104**

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH  
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR  
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI  
IVQENAGSILCSIHAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK  
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF  
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT  
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG  
TPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259



CCTTACACAGGACTCTTCATTGCTGGTTGGCAATGATGATGATATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG  
TTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA  
CTCACTGTTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC  
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCAACAATTTTACAGAAATGAGCCAGAGACTTGAAT  
CAATGGTGAAAAATGCATTTTATAAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC  
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTG  
AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG  
ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA  
ACACGAAGAAGTAAACCTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG  
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC  
TTGTGAGTGTCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA  
ATAAACCTTCGAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACCCCATCACA  
TGACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGT  
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT  
GATGTTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC  
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT  
GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG  
AGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTG  
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTG  
GGTGTGGAGGCCATTTTACAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA  
ATAAACTGTTTTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA  
GATCAACTCTGTCTATCTGTGAGCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATAACAATATTAC  
ATTACAGCCTGTATTCAATTGTTCTCTAGAAGTTTTGTGCAATTTTGACTTGTGACATAAAATTTGTAAT  
GCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTTCAGCAAAATATCCATTT  
TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAAATCCCTACATTTTATTGGCACAGAA  
AAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCAATTATGAAAGGTCAAGCAAGACA  
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG  
ATATATCCTTATTTTCATTTCCAAACAACCTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT  
ATAATAATTATACAAACTTCATGCAATGTACTTGTCTAAGCAAAATTAAGCAAAATATTTATTTAACATTG  
TTACTGAGGATGTCAACAATAACAATAAAATATAAATCACCCA

[illegible]

Transmembrane domain:  
amino acids 21-40 (type II)

**FIGURE 107**

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG  
 CCCCAGATGAGCCCCCGCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCCCAGC  
 GCCGACGATCGCTGCCGTTTTGCCCCTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT  
 ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA  
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTTGATGAGGGACTACCTAA  
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTCCCT  
 GTGGTTCATCCGGTCATGATTGCTGTTTGTGTTTCTTATCATTTGTGGGGATGTTAGGATATTG  
 TGAACGGTGAAAAGAAATCTGTTGCTTCTTGTCATGGTACTTTGGAAGTTTGTGTCATTTTCT  
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA  
 GATATGGTCACTTTGAAAGCCAGGATGACAAATTTATGGATTACCTAGATATCGGTGGCTTACTCA  
 TGCTTGAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG  
 AAATGACAGAGATGGACTGGCCCCCAGATTCCCTGCTGTGTTAGAGAATTCCCAGGATGTTCCAAA  
 CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT  
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA  
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG  
 ACAGACCAAATGATGTCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCCTCAGTAGAACT  
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT  
 TTGAGATGGAGGAGTTATAAAGAAATGTCACAGAAGAAAACCACAACTTGTTTTATTGGACT  
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAATGTTGCCATAAAA  
 TAACACCTAAGCATATACTATTCTATGCTTTAAAAATGAGGATGGAAAAGTTTCATGTCATAAGTC  
 ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC  
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA  
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA  
 CTAGTATATAAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTTATTA  
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACTTTCAATATTGGTGACT  
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACAT  
 TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGCTGTGTAAATCTGTATAATTCAGTCGAT  
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCTGTATAGCATCATT  
 ATTTTTAGCCTTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC  
 TGTTATTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC  
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTTCAGA  
 AAGGACTTGTATGCTGTTTTCTCCCAAATGAAGACTCTTTTTGACACTAAACACTTTTTAAAAA  
 GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA  
 TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT  
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTACAA  
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCAATTTACTGTATTTTGTGTATTTTGTATTAT  
 TTCTCAGAATATGGAAGAAAATTAATGTGTCAATAAATATTTTCTAGAGAGTAA



**FIGURE 109**

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTTCTCCTGATAAAGCCCTACCAGTGCT  
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA  
 CCAGTATTAAGAGGATTTCCAGTGTTCTGGCAGTTGGTCCAGAAGGATGCCCTCCATTCTGCTTCTCACCTG  
 CCTCTTCATCACAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC  
 CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGAG  
 TGGTACCCTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA  
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT  
 TCAATGGGAAGTCTGTCTCTGGAACACCACGGTGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT  
 CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTATGACATCTGCGACGAGGACTGCCATGG  
 CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG  
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTACCGCTGT  
 GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA  
 TGGTGGCTGCAGCCACTTGGCCTTGATCTGAGAAAGGCTACCAGTGTGAATGTCCCGGGGCTGGTGTGT  
 CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG  
 CTGGTTGGTGGCTGGAGCTCTCCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT  
 CCTCTTCTCTCAAGACATGTGGTACAGTGGTCTGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA  
 CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG  
 ACCTGCCAGTTCACGCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAT  
 CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGAGTTGAAGAGCCTTACC  
 GGGAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC  
 TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCACCTCCAAGATCGACGAGGTCCTGAAATACTACCTCAT  
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG  
 TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTTCTTGTCTGTGGAGTG  
 TTGGACGAGCGTTCCCGCTGTGCCCAGGTTGCCACCGGCAATGCGTCTGGGGCAGGAGGAGAGGACTCAGC  
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC  
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTTAAGAACATCTGCCAACAGC  
 TGGGTTCAGACTTCACACTGTGAGTTCAGACTCCAGCACCAACTCACTCTGATTCTGGTCCATTAGTGGGCA  
 CAGGTACACAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTCTAGGGTTGAAAACCTAACTGTCCA  
 CCCAGAAAAGCACTCACCCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGAATCAGAC  
 CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAACAGTTACTGAAATTATGA  
 CTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA  
 TTATAGGGAATTTGGAAGTGTATCAATAAAACAGTATATAATTTT

## FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP  
 TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV  
 YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV  
 LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRGLVLSNHTCQVPVLCKSNAIEVNIIPRELVGG  
 LELFLTNTSCRGVSNGTHVNILFSLKTCGTVDVVDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT  
 CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLRLDSLYFGIEPVVHV  
 SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPEVFKFVGKDHKEVFLHCRV  
 LVCGVLDERSRCAQGGCHRRMRRGAGGEDSAGLQGGQTLTGGPIRIDWED

### Important features of the protein:

#### Signal peptide:

amino acids 1-16

#### N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

#### Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

#### N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
 522-528, 531-537

#### Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

#### ZP domain proteins.

amino acids 431-457

#### Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

## FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG  
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT  
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT  
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC  
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT  
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCCGA  
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC  
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC  
 AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG  
 GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG  
 GAGCTTCGATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGCCTG  
 TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGG  
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCTGGACCCCCACTGGGTCTCAGGCA  
 GCCCCTGCTTCAGGAAACATACCGATGTGTCAACTGGAAGTGGGGCAGGCTCAGACAACTGGGCAG  
 CTTCCTCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA  
 TCGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT  
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA  
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTGATTGACAGCACACGGTGCAATGCAGACGATGCGT  
 ACCAGGGGGGAAGTCACCGAGAAGATGATGTGTGAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT  
 GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG  
 CTGGGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT  
 GGAAGGCTGAGCTGTAATGCTGCTGCCCCCTTGCAGTGTGGGAGCCGCTTCCTTCTGCCCTGCCACCT  
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTGGGTACACCCCTCTGCCCACAGCCTCAGCAT  
 TTCTTGAGCAGCAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGGCGCCAGAGGAAGTCA  
 GCAGCCCTAGCTCGGCCACACTTGGTGTCTCCAGCATCCAGGGAGAGACACAGCCCACTGAACAAGGTCT  
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC  
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA  
 GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT  
 ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA  
 AAAA

[illegible]

Transmembrane domain:  
amino acids 32-53 (typeII)



# FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA  
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTTCAGCAACTAAAAAGCCACAGGAGT  
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAATAATC  
TGTTTTTTGTTCTTTGTAAC TAGCCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT  
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTTCGAAG  
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT  
GGAGGCACTGACTCGGGCAGTGACGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC  
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC  
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGGCGTGGACACCTGTCTCA  
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCGCATCTAACCTTTTCATGTCTGCACATCACCTG  
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC  
CAGAAGGGGTCTGCTTAGACCACCTGGTTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG  
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTGTGCCAAATTATGGGTGAGAAAAGATG  
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG  
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT  
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA  
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC  
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG  
GGCTAATGGCTCAGTGTGGGCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT  
GTGCGAACACGGAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG  
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGTTT  
CAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTGAATTTGAAACCCCAAATCCA  
AACC'TAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT  
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG  
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC  
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA  
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA  
TGTTTATTTGTAA

## **FIGURE 114**

MLWWLVLLLLPTLKS<sup>1</sup>VFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC  
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

**Signal peptide:**

amino acids 1-15

# **FIGURE 115**

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA  
 AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA  
 ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT  
 GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT  
 TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAAGTAAATA  
 TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG  
 CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT  
 CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG  
 ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTAAAAATTCAAAATTCCTGGA  
 GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG  
 ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCTGCCAACGAAAAAAGGGATTGAACAAAAT  
 GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA  
 AGAAGTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG  
 GTTATTGTTGTATTTACTGCCGTGAGGCAACCGCTATTGCCGCGCGTCTGTGAACCTTTACTA  
 GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG  
 TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC  
 TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT  
 GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA  
 TTATCACCAAAAAAAAAAAAAAAAAAAAA

## FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALT LIVLFWGSKHFWPEVPPKKAYDME  
HTFYSGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTG IYFVGLQKCFIKTQIKVIP  
EFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKN SKILEICDNVTMYWINPTLISVSE  
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPM LD  
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

### Important features of the protein:

#### Signal peptide:

amino acids 1-40

#### Transmembrane domain:

amino acids 25-47 (type II)

#### N-glycosylation sites.

amino acids 94-97, 180-183

#### Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

#### N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

#### Microbodies C-terminal targeting signal.

amino acids 315-317

#### Cytochrome c family heme-binding site signature.

amino acids 9-14

## FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA  
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCTCTGTCCATCCTGGGGCT  
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT  
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC  
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCT  
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG  
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG  
TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG  
GATGGTGCAGACTGTTAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGGGGCTGGGTGCTGGAGGCC  
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCACTACAAA  
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT  
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC  
CTTCCAAGCACGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA  
CCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT  
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAATATTCCACCATAAAACA  
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTCAATCCTCTATTTCTTTTTTAAATATAACT  
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC  
TTCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAA  
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA  
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATAATTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT  
ACACTGTGATCTTAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTG  
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCTCTTCTGT  
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCATAATATAGTTAA  
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG  
GAAATGAAAAATAATTGCTTTGACATTGCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCC  
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAAGT  
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA  
TGGTGGCATACACCTGTAGTCCAGCATTCCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAAGGAGGT  
TGGGGCTGCAGTGAGCCATGATCACACCAC'TGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA  
AATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGTAGGTTAAACTAATTCTTTAA

**FIGURE 118**

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP  
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL  
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA  
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY  
V

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

## FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC  
 AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG  
 GTGGTGTGGGATGGTGGGCACAGTGGCTGTCACTGTCACTGCCTCAGTGGAGAGTGTGCGCCTTCATT  
 GAAAACAACATCGTGGTTTTTGAAGCTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA  
 CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACCTACAGGCAGCCAGAG  
 GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC  
 ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGTGACGGCTGGAATCATCTTCAT  
 CATCACGGGCATGGTGGTGTCTCATCCCTGTGAGCTGGGTGCAATGCCATCATCAGAGATTTCTATA  
 ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA  
 CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTGTGCAACGAAAAGAGCAGTAGCTA  
 CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG  
 TCTACTCCAGAAGTCAGTATGTGTAGTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG  
 AAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT  
 AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA  
 TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA  
 CTCTTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC  
 ACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGT  
 TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA  
 ATCATGGATAGGGTTGAAGAAGTTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCTCCCA  
 TTTATAATGAAGATTAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT  
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTCTCTCCAGAGGCTTTTTTTTT  
 CTTGTGTATTAAATTAACATTTTTAAACGCAGATATTTGTCAAGGGGCTTGCATTCAAAGTGCCTT  
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG  
 AAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT  
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA  
 TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT  
 GCTCTATGTGGGTACCTGAGTCAAAATTGTCATTTTGTCTGTGAAAAATAAATTTCTTCTTGTGA  
 CCATTTCTGTTAGTTTTACTAAAATCTGTAAATACTGTATTTTCTGTTTATTCCAAATTTGATGAA  
 ACTGACAATCCAATTTGAAAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT  
 TATACATTTATATTAATAAATTGTACATTTTCTAATT

## **FIGURE 120**

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK  
IYDSLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG  
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY  
RYSIPSHRTTQKSHTGKKSPSVYSRSQYV

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 82-101, 118-145, 164-188



## FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG  
CCAGACGCTGACCACGTTCTCTCTCGGTCTCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC  
GGGAGCCATGCGACCCAGGGCCCCGCCGCTCCCCGAGCGGCTCCGCGGCCTCCTGCTGCTCC  
TGCTGCTGCAGCTGCCCGCGCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG  
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC  
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG  
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACACTAC  
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT  
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTGTTCAGTGCTCACTTCGGCTAAAATGCA  
GAAATGCATGCTGTGACGCTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC  
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG  
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG  
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC  
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC  
TTGGAATGGTTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG  
CTAAATATGTTTACAGACCAAAGTGTGATTTTCACTGTTTTTAAATCTAGCATTATTCATTTTG  
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT  
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTGTTTTTCTCTTAGTATAGCATTTTTTA  
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT  
TAAATAAAAATTATTTCCAACA





## **FIGURE 124**

MGFNVIRLLSGSAVALVIAPT VLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG  
LSLRYSNLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR  
PVTNLRNLDSLQNLHSLGSEQFRGLRKLLSLHLRSNSLRTPVRI FQDCRNLELLDLGYNRIRS  
LARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTSWTWSSLQRLDL  
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIQEILD SWISLNDISLAGNIWECSRNICSLVN  
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE  
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR  
SLMRRHRKKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPTYNKSGSRECEV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 420-442

**N-glycosylation sites.**

amino acids 126-129, 357-360, 496-499, 504-507

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 465-468

**Tyrosine kinase phosphorylation site.**

amino acids 136-142

**N-myristoylation sites.**

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

## FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCTCCGTCCTCCGAGGAGGAGGAGAGGCTTTTGCCGCTG  
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT  
AGCAACCTTTCCCTGGATCTCAGAAAACCTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC  
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC  
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT  
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG  
ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT  
TTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAGATGCAAATGGAAGGAAAAAGGAACTGGAAGG  
AAAACCATTGCGATTTTCGTGGTGACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC  
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA  
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC  
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA  
TCAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG  
ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC  
ATCTTGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA  
TGAGTGGAGTCAGTCCATTTTAA

## **FIGURE 126**

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES  
APYRCGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVKSEDEHYPLWKS  
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP  
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP  
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

### **Transmembrane domains:**

amino acids 25-38, 130-147, 233-248

**FIGURE 127**

CGCGGATCGGACCCAAGCAGGTCGGCGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC  
 CCCCCTGTCTGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG  
 AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCGCTCGGGGGGCATCGGCGCGGCCGTGGCC  
 CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA  
 GCTGGCTGCTGAATGTAAGAGTGACAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTAT  
 CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC  
 TGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAAGTGGTTGGAA  
 GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA  
 AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA  
 CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCAGTGGCTGACAGAGGGACT  
 GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG  
 AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA  
 ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCGCACA  
 CATCCAGATTGGAGACATCCAGATGAGGCCACCGAGCAGGTGACCCTAGTGACTGTGGGAGCTCC  
 TCCTTCCCTCCCCACCTTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT  
 CACGGGATACCACTTCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA  
 TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC  
 TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCTGGGCACCTTGGCCTTTGTCTGCTCTCAGTG  
 TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCGTG  
 TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTATCCAGGGCTCC  
 AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC  
 AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG  
 CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAA  
 AAAAAAAAAA

*(The following information was obtained from the records of the Federal Bureau of Investigation, Department of Justice.)*

Important features of the protein:

amino acids 1-17

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,  
199-205

amino acids 30-42, 104-114





### **FIGURE 130**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV  
KGGHTLGESPMFPKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-myristoylation site.**

amino acids 86-92

**Zinc carboxypeptidases, zinc-binding region 2 signature.**

amino acids 68-79



[illegible]

Important features of the protein:

amino acids 1-28

amino acids 1-28

## amino acids 531-552

amino acids 531-552

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

## amino acids 515-522

amino acids 515-522

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561, 640-645

## amino acids 567-570

amino acids 567-570

## amino acids 159-180

amino acids 159-180

## amino acids 34-44

amino acids 34-44

**FIGURE 133**

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCCAGGCTTCTTGGCA  
 GCCCTGCCGGGCCACTTGTCTTCATGCTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG  
 TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG  
 GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT  
 CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT  
 GCGCTGGAGAGAGCCACCGTCTTCTTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT  
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGGCCAGGAGCCCCCTGC  
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC  
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCCG  
 GTTTTGGAAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCCGGC  
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACGGGGACG  
 GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCCGGCTGCTC  
 AGGCTACTGCCTGTCCCACTGCTCTTCTTCTCTGCGCCAGAATGAGGGGATGCACACAGG  
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTGAACCGC  
 AGAGCTGAGGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG  
 TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA  
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA  
 TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT  
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC  
 AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACCTGGCTAAT  
 TTTTATATTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT  
 CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTTGTAGCCACCGTGTCTG  
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC  
 ATGGGGGCTCTCTCCCCTAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGG  
 GTGGCTTCCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACCCCATCCACACCG  
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC  
 TTTAGTCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG  
 ATAAGCAAAGCCACCCCGACACCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGTTAGGTGGG  
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA

## FIGURE 134

MSARGRWEGGGRACRGSGLGLARAQGAERTVSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD  
 TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV  
 GMLGEKLEAAIQRSIHLYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE  
 RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPIQQSQD  
 YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQKEGCFG  
 EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSLILP  
 SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

### Important features of the protein:

#### Signal peptide:

amino acids 1-26

#### Transmembrane domain:

amino acids 39-56

#### Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

#### N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

#### Amidation site.

amino acids 10-14

#### Glycoprotein hormones beta chain signature 1.

amino acids 230-237

# **FIGURE 135**

GGTCTGAGTGCAGAGCTGCTGTCTATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCTGCTGCTG  
CTGCTGCTATCGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG  
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAATTGTTCCAGGGGTGAAGC  
CTCAGGACTGGATCTCGCGGCCCGAGTGTGGTAGACGGAGAAGAGCACGTCCGGTTTCTTAAG  
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC  
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG  
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT  
CCACCTTCTTACTTTATTAAAAGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT  
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC  
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT  
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG  
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGAGTCAGGCCGTCCAGAGCTGGCATTTCAC  
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAACCGTGTGAAGCAACTACTATAAACTT  
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT  
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT  
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAATTAT  
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAA

## **FIGURE 136**

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEAGSGGSGVGIGDRFKIEGRAVVPGVKPDWISAA  
 RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVS PAYRFDPVRVDITSKGKMRARYVNYIKTSE  
 VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMMLVPLLI FVLLPKVVNTSDPDMRREME  
 QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

### **Important features of the protein:**

#### **Signal sequence:**

amino acids 1-23

#### **Transmembrane domain:**

amino acids 161-182

#### **N-glycosylation site.**

amino acids 184-187

#### **Glycosaminoglycan attachment sites.**

amino acids 37-40, 236-239

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 151-154

#### **N-myristoylation sites.**

amino acids 33-38, 36-41, 38-44, 229-234

#### **Amidation site.**

amino acids 238-241

#### **ATP/GTP-binding site motif A (P-loop).**

amino acids 229-236



## FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA  
 CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA  
 GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC  
 CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC  
 CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT  
 TTTTCTTACTCCTGTGGTCTCCAGATTTAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC  
 CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG  
 AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG  
 TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTGAAGACACA  
 AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTTAA  
 AAACCTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA  
 CCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACATGCCATTGTGGGGAGGAAGCAATG  
 AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC  
 TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG  
 CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA  
 ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG  
 CTTCTTGCATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT  
 TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTATTTA  
 ATGTATTTATTTTTTACTTGGACATGAAACTTTAAAAAAATTCACAGATTATATTTATAACCTG  
 ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGG  
 CTAGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT  
 ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTGTATGTGAATTGCAC  
 ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG  
 CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAAA  
 AAAAAAAAAA

## **FIGURE 138**

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF  
 SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG  
 FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS  
 SLANSFLTICKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET  
 E

**Important features of the protein:**

**Signal peptide:**

amino acids 1-42

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 192-195, 225-228

**N-myristoylation sites.**

amino acids 42-47, 46-51, 136-141

## FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC  
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC  
TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTTGCCAGAGCGGAA  
CACGGAGCGGAGCCCCAGCGCCGAACCTCGGCTGGAGCCAGTCTAACTGGACCACGCTGCC  
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAGATGATTTTGTGTTG  
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT  
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG  
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA  
TAAGACAAGCAAAGCAGCTGGAAAAAGAAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA  
TTCCCAAAGCATTGTTGTGAATATGATGGAACAACCTCATTTTCGTGTCAGATTACACAGAGGT  
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAGAGAGAGATTGGGACATATTAATCC  
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTACAGGGCCCAACAGCCCCCTGATTGGGCAG  
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA  
GACGCCCTTACCCAATTGCTGGTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG  
GGGCCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTAATCAGTTCGCGTTTGAAGGAAA  
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC  
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGAC  
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG  
AATGTGCCGTCATATGAAAAAGATCCTGGGTTTGAAGCAGTTTAAATGTCAGAAAGATTGCATGG  
GAAGTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAAGTCTATTCAACCTGGGCTCCAAGG  
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGACAAAGTGCCCCAG  
TTCTCACCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA  
CTGTCTCTCCTGGGTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTTCGGCCGTTACGT  
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCTCTCGTGGCTGGCGGCAGGCT  
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC  
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCGACACAGGTG  
TTCACATCTGTGCTGTGAGGTGAGATGCCTCAGTTCTTGGAAAGCTAGGTTCTTGCGACTGTAC  
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA  
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA  
CACTCAGGAGGGTCAAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT  
GCCTCATCAGGTCCAGATTTCTTTCAAGGCGGACGTTTTCTGTGGAATTCTTAGTCTTTGGCC  
TCGGACACCTTCATTGCTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC  
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC  
ACCCCAACCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTGAGTGT  
CTGACCGAGACACTCACAGCTTTGTCTATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG  
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG  
CTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

[illegible]

Important features of the protein:

amino acids 1-34

## amino acids 58-76

## amino acids 56-60, 194-198

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,  
275-281, 278-284

## amino acids 154-158

## amino acids 205-208

## FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA  
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC  
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC  
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT  
AGCAGTTCCAGATAAAAACTACATACGCCAGAGATCTTCTTGCATTAGCCTCATCCTTGAGCT  
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC  
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT  
GGCTGCCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA  
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT  
GTTGGGGTGACAGATAAAATTTGAGAACAGGAAACACATTGAATTTTCATTTC AACAGTTTGCAA  
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA  
ATTTGAACTAATTGTATAAAAAACCAAACCTGCTCACT

## **FIGURE 142**

MLLLLLLEYNFPIENNCQHLKTTHTFRVKLNLPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI  
FFALASSLSASAEGKSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI  
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 33-36

**N-myristoylation site.**

amino acids 50-55, 87-92

**Interleukin-1**

amino acids 37-182

**FIGURE 143**

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA  
TCCAGGATCCTGTCCTTCCTGTCCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG  
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTCTGTGGG  
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG  
AAGGGAGGTCTGTCCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC  
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCCTGCCA  
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG  
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT  
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACTGCCATTTCCAAGAAAGCACAG  
AGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC  
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTGAGTGAAACCCACTCACAGGCTTGTCCATGT  
GCTGCTCCACATTCCTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT  
TTGGACTTGTGTTTATCCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTGTTTAGAAAATCC  
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAAAA  
AA

## **FIGURE 144**

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY  
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP  
WMTQFSLLNKTCLEGFH

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-glycosylation sites.**

amino acids 117-121, 139-143

**N-myristoylation site.**

amino acids 9-15



**FIGURE 145**

CTGTGTCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAAGTTCTGACGCGATGAGGAAG  
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA  
GACGAGGGGCATCAAGCACAGAATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCCAGATCA  
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT  
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCTCCGATGGCATCCA  
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCAACGGCTGCATCAATGCCA  
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG  
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT  
TCGGGTACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT  
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT  
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA  
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT  
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT  
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG  
GTGTATTCCCGTTTTACAGATAAGGAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC  
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT  
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAACATGTAT  
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

## **FIGURE 146**

MRKHLSWWLLATVCMLLFSLHSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK  
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ  
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVILLCLLALIWLMLVK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 157-171

#### **N-glycosylation sites.**

amino acids 98-102, 110-114

#### **Tyrosine kinase phosphorylation site.**

amino acids 76-83

#### **N-myristoylation sites.**

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

#### **Amidation site.**

amino acids 62-66

# **FIGURE 147**

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT  
CCTGATGATTATAGACTCAAAGAAAACATGTTCAGAAGCTCTTCTCTTCTGGCCTCCTCT  
CTGTCTTCTTCCCTCTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG  
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTTGAAATTT  
CAACTTTCAGATTCAGGGGTACATGTGAAGGTTGTTTTATGAGTATATTGCATTGATGTGCTGAGG  
TTTGGGGT

## **FIGURE 148**

MFRSSLLFWPPLCLLSLFLLLISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE  
GLFYEYIA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-myristoylation site.**

amino acids 62-68

## FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC  
TCCGCGCGCGCCCAACCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCCCAGTCC  
CGGCCCCCTCTCCGCGCCACACCCACCCTCCTGGCTCTTCTGTTTTTACTCCTCCTTTTCATTTCATA  
ACAAAAGCTACAGCTCCAGGAGCCAGCGCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT  
CCTCGGGACCGGCACTTGGATTCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA  
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA  
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT  
TGTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA  
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG  
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA  
TCAACTAGACGGGACTCCTTTAACCCTGGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG  
AAAATGACAGAGCCGTGTTTGACAAGATTGTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC  
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA  
TTATGAGGAGGATCCCAATAAGCCACAAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG  
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA  
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA  
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAAGAAAGCAAAAGAGAAAGAAA  
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT  
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT  
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA  
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGAAAAGGAATATGGAAGCTTGAAGGATTCACAAAA  
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC  
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGAAATAAAGAAGATTATGACCTTT  
CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA  
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA  
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATCTGTGATTAAAAATTTTTGACCCAAGG  
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC  
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAGAAAAAAGAAAAA

## **FIGURE 150**

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG  
 QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK  
 FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE  
 VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMAAIQDGLAKGENDETVSNTLTLTNGLE  
 RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV  
 SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD  
 NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK  
 EEAEAIKRIYSSL

### **N-glycosylation sites:**

amino acids 68-71, 346-349, 350-353

### **Casein kinase II phosphorylation site:**

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-  
 220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-  
 341, 369-372, 382-385, 386-389, 387-390

### **N-myristoylation sites:**

amino acids 143-148, 239-244

**FIGURE 151**

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA  
GATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC  
 ATAATAACCAAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC  
 GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCTGGGTGTCCAGGGTGGAAAG  
 CCAGTGCCTGTCTGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG  
 AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC  
 TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC  
 TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC  
 AGCAGTGTGACTAGGGCAACGTGCCCCCAGAATCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT  
 GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG  
 TGGGCACCTGACCACTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTTGGAGCTCAGT  
 CCACGGTCCCTCCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAAA  
 CTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT  
 TAATGGTAAC TGACAAGTGTTACCCCTGAGCCCCGAGGCCAACCATCCCCAGTTGAGCCTTATA  
 GGGTCAGTAGCTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTCATA  
 GAGTCAGGGATCTATGGCCCTTGGCCCCAGCCCCACCCCTTCCCTTAAATCCTGCCACTGTCTATA  
 TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGATGTGAGAA  
 GAAATGGCTCGAGCTCAGAAGATAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA  
 GATACATCAAAATCCCAGATGCTGGTCTCTATTTCCCATGAAAAAGTGTCTATGACATATTGAGA  
 AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATTAAGATACCTATTTATATATT  
 TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT  
 AGGTGATTTTCTTTTAAATCTGTTAATTTATCTGTATTTCCCTAATTTTCTACAATGAAGATGA  
 ATTCCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA  
 TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG  
 TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT  
 GTGGCTGGAATCTCTGGGTAAGGAACTTAAAGAACAAAAATCATCTGGTAATTTCTTCTAGAAAG  
 GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA  
 ATTTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTCAGTCTGTGAGTTTATTTGGAGATAAG  
 GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCATAT  
 GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA  
 AAGATGAAGGCAGAGATCGGAGTTTTTGACGCCACAAGCTAAGAAACACCAAGGATGTGGCAACC  
 ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATTTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG  
 CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGATAAATTTTCGGCTGTTTTAA  
 GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT  
 CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTGTGACCAA  
 TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC  
 TACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG  
 AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT  
 GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT  
 AAGTTGCTCAGTTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

## **FIGURE 152**

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESVVPNRWLDASLSPVILGVQGS  
 QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP  
 VRLTQLPENGGWNAPITDFYFQQCD

### **N-myristoylation sites.**

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

### **Interleukin-1 signature.**

amino acids 111-131

### **Interleukin-1 proteins.**

amino acids 8-29, 83-120, 95-134, 64-103



**FIGURE 153**

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC  
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG  
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC  
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA  
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT  
ATCTGATGAAGCAGGTGCTGAACCTCACCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTC  
CAGCCTTATATGCAGGAGGTGGTGCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA  
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC  
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT  
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA  
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG  
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA  
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTCAATTGTAAGTGGTGTTT  
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTTTCCATAAAAAAGATTACTTTCCAT  
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA  
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT  
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT  
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

**FIGURE 154**

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL  
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS  
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

**Important features of the protein:**

**Signal peptide:**

amino acids 1-33

**N-glycosylation sites.**

amino acids 54-58, 68-72, 97-101

**N-myristoylation sites.**

amino acids 14-20, 82-88

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 10-21

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCGCTGT  
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCAGTGCAGAGGGC  
TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC  
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT  
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT  
TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAAGGG  
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG  
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGGCCATCT  
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCGT  
TGCTGTGCCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCCGGGGCAACTCGGA  
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA  
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCCTTAGCTTGTGTGTGTGCGGCCCCGT  
GTGATGGGCTTAGCCGACCTGCTGGAGGCTGGTCCCTTTTTGGGAAACCTGGAGCCAGGTGTACA  
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCGTGAAGTGTGTCTGGAGCAG  
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACCTCTGCACATTTTGAAAAGAG  
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCACTTTTCTCTCAGGAAAGGTTTTCAAA  
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG  
GCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT  
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGATGTTACTAGTCTT  
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT  
CTTTATTTAAAAATGAAAAA

## **FIGURE 156**

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCKGQDTSEELLRWSTVPVPPLEPA  
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS  
ELLYHNQTVFYRRPCHGEKGTGKGYCLERRLYRVSLACVCVRPRVMG

**Important features of the protein:**

**Signal peptide:**

amino acids 1-32

**N-glycosylation site.**

amino acids 136-140

**Tyrosine kinase phosphorylation site.**

amino acids 127-135

**N-myristoylation sites.**

amino acids 44-50, 150-156

## FIGURE 157

CCGGCGATGTCGCTCGTGCCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCGAGAGCC  
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC  
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAGTAGTGTGCAACAGGGGACTATTCA  
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA  
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG  
CCTTCCAGACTCAGACCAGACCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCTGTGA  
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG  
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA  
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGATGAGGAGACA  
GTAGAAGTGAACCTCACAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC  
TATCATCGGGTTTTCTCAGGTGTTTGTAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA  
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCCTACTTGTGGC  
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT  
GGATAACAACAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA  
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT  
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA  
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTCATCCTTGAAA  
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA  
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA  
GAGCGAGGGCAGTCCCAGTGAGAAGTCTCAAGACCTCTTCCCCCTTGCTTTAACCTTTTCTGCA  
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA  
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT  
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG  
ATGGCTGCTGCTCCTTGTAG

**FIGURE 158**

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLDLRDLRVEPVTTTSVATGDYSILMNVS  
 WVRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP  
 NANMNEDGPSMSVNFSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNYMALIQH  
 STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCFQGTGVFPPLDNNK  
 SKPGGWLPLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL  
 QNHRSEVILEKWQKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA  
 FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD  
 GCCSL

**Important features of the protein:****Signal peptide:**

amino acids 1-14

**Transmembrane domain:**

amino acids 290-309

**N-glycosylation sites.**

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283  
 - 287

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 228 - 232 and 319 - 323

**Casein kinase II phosphorylation sites.**

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

**N-myristoylation site.**

amino acids 116-122

**Amidation site.**

amino acids 488-452

## **FIGURE 159**

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT  
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA  
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC  
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG  
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA  
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA  
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT  
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC  
ACTCAGCTGAAGAAG

## **FIGURE 16o**

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN  
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV  
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

**Signal sequence:**

amino acids 1-30

**N-glycosylation site.**

amino acids 83-87

**N-myristoylation sites.**

amino acids 106-111, 136-141



**FIGURE 161**

A C A C T G G C C A A A C A A A A C G A A A G C A C T C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G  
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C C T C T G G A G G C T G A A G A G G G A T T C  
 C A G C C C C T G C C A C C C A C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C C T T G G G G G G G G G C A G C A C  
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C T A G A A G A T G C C T G T G C C C T G G T T C T T G C T G T C C T  
 T G G C A C T G G G C C A A G C C C A G T G G T C C T T T C T C T G G A G A G G C T T G T G G G G C C T C A G G A C G C T A C C  
 C A C T G C T C T C C G G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T  
 C G T G C C T G C T C C G G G C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A G C T G G T G C T G A G G T G C C  
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G  
 G A A G A G C C T G A A G A T G A G G A A A G T T T G G A G G A G C A G C T G A C T C A G G G G T G G A G G A G C C T A G G A A  
 T G C C T C T C T C C A G G C C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C  
 T G G A G G T G C A A G T G C C T G C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C  
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A A A  
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G T C A G C A G A T G G T G A C A  
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G  
 G T C C A G G G C C C C C A A A A C C C G G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A  
 C C A C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A  
 C G A A C A T C T G C C C C T T C A G G G A G G A C C C C G C G C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G  
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G C T G C C C G C A G A A G C G G C A C T  
 G T G C T G G C G G G C T C C G G G T G G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C C T G G G A G A A C G  
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G  
 A A C A G C T C G G A G A A G C T G C A G C T G C A G G A G T G C T T G T G G G C T G A C T C C C T G G G C C T C T C A A G A  
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C C T C T G T G C C T T G G A C C C A  
 G T G G C T G T A C T T C A C T A C C C A G C A A A G C C T C C A C A G A G G G C A G C T C G C C T T G G A G A G T A C T T A C T A  
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G A G C G C A T G G G C C T G  
 C C C C A T G G A C A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G  
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A  
 C A G G A C G T C C G C T C G G G G G C G G C C G C A G G G G C C G C G G C T C T G C T C C T C T A C T C A G C C G A T G A  
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G C T G C G C T G G  
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G C C C G T G G C T T G G T T C A C G C G C A G  
 C G G C G C C A G A C C C T G C A G G A G G G C G C G T G G T G G T C T T G C T C T C T C C C G G T G C G G T G G C G C T  
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C G G G G C G C A C G G C C G C A C G A C G C C T T C C  
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G G C A G C T A C G T G G G G  
 G C C T G C T T C G A C A G G C T G C T C C A C C C G G A C G C C G T A C C C G C C C T T T T C C G C A C C G T G C C C G T C T T  
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C C G C G T T C C G  
 G G C G G C T C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T C A G C C A G C C C T G G A T A G C T A C T T C  
 C A T C C C C C G G G G A C T C C C G C G C C G G A C G C G G G T G G G A C A G G G C G G G A C C T G G G G C G G G G A  
 C G G G A C T T A A A T A A A G G C A G A C G C T G T T T T C T A A A A A A

## FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCRLWDSIDLCLPGDIVPAPGPVLAPTHLQTELV  
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLVSFQAYPTARCVLLEV  
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSTQPRYEKELNHTQQLPALPWLNVSDGDNVHLVLNV  
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWLEPDSVRTNICPFREDPRAHQ  
LWQAARLRLTLQSWLLDAPCSLPAAEALCWRAPGGDPCQPLVPPLSWENVTVDKVFLLKGFHNLQVQ  
VNSSEKLQLECLWADSLGPLKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS  
GQCLQLWDDDLGALWACPMKY IHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG  
RAALLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFHQAQRRTIQEGGVVLLFSP  
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHHPDAVPALFRTVPVFT  
LPSQLPDFLGALQQPRAPRSGLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

### Signal sequence:

amino acids 1-20

### Transmembrane domain.

amino acids 453-475

### N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,  
334-337, 357-360, 391-394

### Glycosaminoglycan attachment site.

amino acids 583-586

### cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

### N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,  
692-697, 696-701, 700-705

**FIGURE 163**

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT  
 GCTCACGCCCCTGAGGACCCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA  
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA  
 AGACGTACGGAGAGAGGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCTGTC  
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT  
 GTCAGTGGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTACGCTCTCTGCAGCACACTAC  
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA  
 CCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGGAAGACATCTTCCATGACCTG  
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA  
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA  
 CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC  
 TACTCCTTCTCCGGAGCCTTCCGTGTTCCCATGGGCTTCCCTCGTCGCAGTACTCTGTACCTGAG  
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTTGACTT  
 TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGCCCCAGC  
 AGTCTGGCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCCGAGGAGC  
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC  
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCCCTGAG  
 GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATTCTACGCCCCACA  
 GGCCATCTCTAAGGTCCAGCCTTCCCTCCTATGCCCCCAAGCCACTCCGGACAGCTGGCCTCCCT  
 CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCTTAA  
 CACCTTAGGCCTAAAGGTGAGCTTTCAGAAAGAGCCACAGCTGGAAGCTGCATGTTAGGTGGCCT  
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC  
 CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA  
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC  
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC  
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTG  
 GAGCAGCCACAGAACTGGATTCTCTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCCCTG  
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA  
 ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC  
 AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCCTCACCAGGAACAAGCAGCATGATA  
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG  
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCAGGTCCAGAGCCCCACCTC  
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCCTCTCCTTGGCCCCATTCTGGCCAGTTTC  
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA  
 GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC  
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA  
 TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTAGCTTCATTCTCTGATAGAACAAAGC  
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCT  
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAACACAC  
 TGTACTGATGTCACAACCTTGCAAGCTCTGCCTTGGGTTAGCCCATCTGGGCTCAAATTCAGC  
 CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCCTC  
 ATCTGTAATGTGGGGATCATAACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG  
 TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAACGGTAGCTATTTAAAAA  
 AAAAAAA

## FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW  
 VAKKGCQRITRKSCNLTVETGNLTLEYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCTIS  
 KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGKGKQREYEFFGLTPDT  
 EFLGTIMICVPTWAKESAPYMCRVKTLDPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP  
 PNSLVQQRVLTFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT  
 YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFFYAPQAISKVQPSSY  
 APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSIAM  
 EESQEAKSLHQPLGICTDRTSDPNVLHSGEETPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC  
 SPSDQGPSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTVQWES

### Signal sequence.

amino acids 1-17

### Transmembrane domain.

amino acids 233-250

### N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

### N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

**FIGURE 165**

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATGG  
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG  
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT  
GTAAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA  
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTGGGCTGGA  
AGTGTGGAACACAGTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA  
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTGAAGGAGGAAGAGATGATT  
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG  
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA  
ACTTGACCCTGTGCCTGAGCCCGAGGCATTGAGAGCTGATTGAGAGGATGGAGAAGGTGCTTTCT  
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT  
CCTGCGGCTAACGCTCAGGGAGTGCACTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA  
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA  
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCGTTATT  
CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTTTTACAAAGATTGTTTTTTAGTA  
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTCCCTTCTAAGTA  
AAAAAAAAAAAAAAAAAAAA

## **FIGURE 166**

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECSMMLMYRGKALEDFTGPDRCFVN  
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFEGGRD  
DFNSYNVEELLGSLELEDSVPESKKAEEVSQHREKSPESRGRELDPVPEPEAFRAEDGEDGA  
FSESTEGLGQGPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGESERTGNSSPASVER  
EKTDAYKVLKTEMSQRGSGQCVIHYSGFRWHQNLSLFYKDCF

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**N-glycosylation site.**

amino acids 294-298

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-34

**Tyrosine kinase phosphorylation site.**

amino acids 67-76

**N-myristoylation sites.**

amino acids 205-211, 225-231, 277-283

**Amidation site.**

amino acids 28-32

**FIGURE 167**

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCATTGTGTCAGAGGCCGGGGAAGAGAAGCAAAGCGC  
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCTCTAGGACATACACGGGACCCCTAACTTC  
 AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGGGCACAGG  
 CGCGGCAGGCGGCAGTCCCGGCCGAAGGCGATGCGCGCAGGGGTCGGGCAGCTGGGCTCGGGC  
 GCGGGAGTAGGGCCCGGCAGGGAGGCGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG  
 GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCGCGCATGAGCCGCGTGG  
 TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCGCGCGCTGGTCAGC  
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACT  
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCTCTCA  
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG  
 ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTGGAGGAATGGAGATGGGCAACATCTGG  
 TGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCACAGTACCGAACTGGTACACAG  
 ATGAACCTTCTGCGGAAGTGAAAGTGTGTTGTGATGTATCACAACCAACTGCCAATCTCTGGC  
 CTTGGGGGTCCCTACCTTTACCAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTATATTTG  
 CAAGTATGAACCAGAGATTAATCCAACAGCCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG  
 GAGACACCCATCAGAATGTGGTTGTACTGAAGCAGGTATAATTCCTAATCTAATTTATGTTGTT  
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTCCAGATGCT  
 GCATAAAGTAAAGGAAGAACAACAACTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA  
 CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTCCAGAAATTTGTAATCT  
 GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGGAAATGGCTTGAAATCACAAGGATCTGC  
 AAGATGAAGTGAAGCTCCCCCTTGAGGCATAATATAAGTAATTTTATATGTCTATTATTCA  
 TTTAAAGAAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA  
 ACTTCAAACCTTCAAGCAATGAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA  
 TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAA  
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAA  
 ATGAAGTGTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG  
 AAACCTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT  
 TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA  
 AATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT  
 TGTCTAATTTCAATTTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT  
 TTGTAATAATGTAACCTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAG  
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA  
 GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG  
 ATTGACACTGGAGGCAGATAGTTGCAAGTGTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG  
 ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG  
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAATGGGTTGG  
 AACCCATCAGTGATCGCATATTCATTGATGAGGTTTGGCTTGAGATAGAAAATGGTGGCTCCTTT  
 CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCCTGTTCTTCTCAAGAGAAAGTTGTAACCTCT  
 CTGGTCTTCATATGTCCCTGTGCTCCTTTAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA  
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 168**

MSRVV\$LLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYKMAFHELSSRVSFQEARLACESE  
GGVLLSLENEAEQKLIESMLQNLTKEPTGISDGFHWGLWRNGDGTSGACPDLYQWSDGSNSQ  
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK  
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKTSPNQ  
STLWISKSTRKESGMEV

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-21

#### **Transmembrane domain:**

amino acids 214-235

#### **N-glycosylation sites.**

amino acids 86-89, 255-258

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 266-269

#### **N-myristoylation sites.**

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217